



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 117161**

**TO: Nancy Vogel**  
**Location: rem/2a65/2c70**  
**Art Unit: 1636**  
**Monday, March 29, 2004**  
**Case Serial Number: 09/963285**

**From: Paul Schulwitz**  
**Location: Biotech-Chem Library**  
**REM-1A65**  
**Phone: (571)272-2527**

**paul.schulwitz@uspto.gov**

### **Search Notes**

Examiner Vogel,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz  
Technical Information Specialist  
STIC Biotech/Chem Library  
(571)272-2527

FEATURES  
source  
Location/Qualifiers  
1..19  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
Db 14 GGAAGTAAAAA 3

## RESULT 2

ARI47023/c ARI47023 28 bp DNA linear PAT 08-AUG-2001  
LOCUS  
DEFINITION Sequence 84 from patent US 6221361.  
ACCESSION ARI47023  
VERSION ARI47023.1 GI:15110826  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
UNCLASSIFIED.

REFERENCE 1 (bases 1 to 28)  
AUTHORS Cochran, M.D. and Junker, D.E.  
TITLE Recombinant swinepox virus  
JOURNAL Patent: US 6221361-A 84 24-APR-2001;  
LOCATION/Qualifiers  
1..28  
/organism="unknown"  
/mol\_type="unassigned DNA"

## FEATURES

source

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 28;  
Best Local Similarity 100.0%; Pred. No. 5.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
Db 27 GGAAGTAAAAA 16

## RESULT 3

B0273541/c B0273541 28 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Recombinant virus expressing foreign DNA encoding feline CD80,  
feline CD86, feline CTLA-4 or feline interferon-gamma and uses  
thereof.  
ACCESSION B0273541.1 GI:33083309  
VERSION JP 2002513581-A/75.  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 28)  
AUTHORS Winslow, B.J. and Cochran, M.D.  
TITLE Recombinant virus expressing foreign DNA encoding feline CD80,  
feline CD86, feline CTLA-4 or feline interferon-gamma and uses  
Patent: JP 2002513581-A 75 14-MAY-2002;  
JOURNAL SCHERING-PLOUGH LTD  
COMMENT OS Unidentified  
PN JP 2002513581-A/75  
PD 14-MAY-2002  
PF 30-APR-1999 JP 2000547248  
PI 01-MAY-1998 US 09/071711  
PT BARBARA J WINSLOW, MARK D COCHRAN  
PC C12N15/09, A61K39/12, A61K39/15, A61K39/215, A61K39/23,  
A61K39/245,  
PC A61K48/00, A61P43/00, C12N7/00, C12N15/00 CC FIV  
PPR downstream primer  
FH Key Location/Qualifiers

FT source 1..28  
/organism="Unidentified".  
FT Location/Qualifiers  
1..28  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 28;  
Best Local Similarity 100.0%; Pred. No. 5.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
Db 27 GGAAGTAAAAA 16

## RESULT 4

AR267926/c AR267926 28 bp DNA linear PAT 10-APR-2003  
LOCUS  
DEFINITION Sequence 265 from patent US 6497882.  
ACCESSION AR267926  
VERSION AR267926.1 GI:29698051  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
UNCLASSIFIED.

REFERENCE 1 (bases 1 to 28)  
AUTHORS Cochran, M.D. and Junker, D.E.  
TITLE Recombinant swinepox virus  
JOURNAL Patent: US 6497882-A 265 24-DEC-2002;  
LOCATION/Qualifiers  
1..28  
/organism="unknown"  
/mol\_type="genomic DNA"

## FEATURES

source

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 28;  
Best Local Similarity 100.0%; Pred. No. 5.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
Db 27 GGAAGTAAAAA 16

## RESULT 5

AX093873 AX093873 58 bp DNA linear PAT 30-MAR-2001  
LOCUS  
DEFINITION Sequence 2 from Patent WO0118039.  
ACCESSION AX093873  
VERSION AX093873.1 GI:13510091  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 58)  
AUTHORS Pitha-Rowe, P.M., Yeow, W.S. and Au, W.C.  
TITLE Increased production of interferon- $\gamma$ (a)  
Patent: WO 0118039-A 2 15-MAR-2001;  
JOURNAL GLAXO GROUP LIMITED (GB)  
COMMENT OS Unidentified  
PN JP 2002513581-A/75  
PD 14-MAY-2002  
PF 30-APR-1999 JP 2000547248  
PI 01-MAY-1998 US 09/071711  
PT BARBARA J WINSLOW, MARK D COCHRAN  
PC C12N15/09, A61K39/12, A61K39/15, A61K39/215, A61K39/23,  
A61K39/245,  
PC A61K48/00, A61P43/00, C12N7/00, C12N15/00 CC FIV  
PPR downstream primer  
FH Key Location/Qualifiers

Query Match 100.0%; Score 12; DB 6; Length 58;  
Best Local Similarity 100.0%; Pred. No. 5e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
Db 27 GGAAGTAAAAA 16

```

QY      1  GGAAGTAAAAA 12
      |||||
Db      12  GGAAGTAAAAA 23

RESULT 6
LOCUS   AF227460/c
DEFINITION Bromus tectorum clone 21 microsatellite sequence.
ACCESSION AF227460
VERSION   AF227460.1 GI:8489262
SOURCE    Bromus tectorum
ORGANISM   Bromus tectorum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Bromaeae; Bromus.
REFERENCE
AUTHORS   Paulsen,A., Meyer,S.E., Coleman,C.E. and Fairbanks,D.J.
TITLE     Microsatellite markers for Bromus tectorum (cheatgrass)
JOURNAL   Unpublished
AUTHORS   Paulsen,A., Meyer,S.E., Coleman,C.E. and Fairbanks,D.J.
TITLE     Direct Submission
JOURNAL   Submitted (21-JAN-2000) Botany and Range Science, Brigham Young
University, 401 WIDB, Provo, UT 84602, USA
FEATURES
source    1..150
            /organism="Bromus tectorum"
            /mol_type="genomic DNA"
            /db_xref="taxon:29667"
            /clone="21"
repeat_region 1..150
            /note="microsatellite"
            /rpt_type=tandem

ORIGIN
Query Match 100.0%; Score 12; DB 8; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGAAGTAAAAA 12
      |||||
Db      58  GGAAGTAAAAA 47

RESULT 7
LOCUS   AX305717
DEFINITION Sequence 468 from Patent WO0188188.
ACCESSION AX305717
VERSION   AX305717.1 GI:17645138
KEYWORDS Mus musculus (house mouse)
SOURCE    Mus musculus
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS   Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE     Method for examining ischemic conditions
JOURNAL   Patent: WO 0188188-A 468 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source    1..151
            /organism="Mus musculus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:10090"

ORIGIN
Query Match 100.0%; Score 12; DB 6; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  GGAAGTAAAAA 12
      |||||
Db      102  GGAAGTAAAAA 113

RESULT 8
LOCUS   BX322143/c
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.35925, sequence tagged site.
ACCESSION BX322143
VERSION   BX322143.1 GI:29420274
KEYWORDS STS; STS, sequence tagged site.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS   Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
TITLE     Unpublished
JOURNAL   2 (bases 1 to 191)
AUTHORS   Clarke,J.H.
TITLE     Direct Submission
JOURNAL   Submitted (31-MAR-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT    At denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, Gr a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3' end of the
transposon. 5 denotes a sequence derived from the 5' end of the
transposon. BSRC GARNET, Afis project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N122636.
FEATURES
source    1..191
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /variety="Columbia-0 NASC stock code N1092"
            /db_xref="taxon:3702"
            /clone="AC006841"
            /note="Derived from superpool 25.06 NASC code N41149"

STSS
1..191
/standard_name="SM_3.35925"

ORIGIN
Query Match 100.0%; Score 12; DB 11; Length 191;
Best Local Similarity 100.0%; Pred. No. 4.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGAAGTAAAAA 12
      |||||
Db      89  GGAAGTAAAAA 78

RESULT 9
LOCUS   G55264/c
DEFINITION SHGC-100532 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G55264
VERSION   G55264.1 GI:6120583
KEYWORDS STS.
SOURCE    Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS   1 (bases 1 to 206)
TITLE     Olivier,M. and Cox,D.R.
JOURNAL   Unpublished, Olivier, M., Cox, D.R. 2000
Unpublished (2000)

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 08:39:03 ; Search time 6718.14 Seconds  
(without alignments)  
6361.316 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1250\_2235

Perfect score: 986

Sequence: 1 ctgcattccatccagcgc.....gagccgtctcggaagcagca 986

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.red.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	986	100.0	3289	9	HSMPFH1	Y08223 H.sapiens M
2	954	96.8	168656	9	AC009108	AC009108 Homo sapi
3	298.6	30.3	178416	10	AC127554	AC127554 Mus muscu
4	290	29.4	6021	10	MMMFHEAD1	Y08222 M.musculus
5	88.4	9.0	159805	2	BX546449	BX546449 Danio rer
6	83	8.4	869	11	PM2A12B	AL684671 Penicilli
7	80.6	8.2	172457	2	BX664728	BX664728 Danio rer
8	79.8	8.1	2000	6	AX655393	AX655393 Sequence
9	79.8	8.1	51753	2	AC124247	AC124247 Homo sapi
10	79.8	8.1	75144	2	AC027483	AC027483 Homo sapi
11	79.6	8.1	885	11	PM7F8G	AL685189 Penicilli
12	79.6	8.1	11305	8	OS4535061	AJ535061 Oryza sat
13	79.2	8.0	85434	2	AC066610	AC066610 Homo sapi
14	79	8.0	224777	2	AC138109	AC138109 Mus muscu
15	78.8	8.0	93821	2	AC021596	AC021596 Homo sapi
16	78.6	8.0	177948	2	AC146089	AC146089 Pan trogl
17	78	7.9	300695	2	AC079431	AC079431 Mus muscu
18	77.6	7.9	192042	2	AC063969	AC063969 Mus muscu
19	77.4	7.8	1279	11	PM2H12G	AL684840 Penicilli
20	77.4	7.8	181850	2	BX276102	BX276102 Danio rer
21	77.4	7.8	21952	2	AC084804	AC084804 Mus muscu
22	77	7.8	63540	2	AC110231	AC110231 Mus muscu
23	77	7.8	69570	2	AC067888	AC067888 Homo sapi
24	76.8	7.8	60931	2	AC105113	AC105113 Homo sapi
25	76.2	7.7	205691	2	AC087227	AC087227 Mus muscu
26	76	7.7	207420	2	AC078884	AC078884 Mus muscu
27	75.6	7.7	69570	2	AC067888	AC067888 Homo sapi
28	75.2	7.6	81767	2	AC021829	AC021829 Homo sapi
29	75.2	7.6	155657	2	BX469901	BX469901 Danio rer
30	75.2	7.6	216645	2	AC119156	AC119156 Mus muscu
31	75	7.6	135552	2	AC019251	AC019251 Homo sapi
32	75	7.6	190822	2	AC051613	AC051613 Mus muscu
33	75	7.6	216332	2	AC145342	AC145342 Pan trogl
34	74.8	7.6	78220	2	AC023212	AC023212 Homo sapi
35	74.8	7.6	138709	2	BX649540	BX649540 Danio rer
36	74.8	7.6	217412	2	AC024400	AC024400 Homo sapi
37	74.6	7.6	52884	2	AC068018	AC068018 Homo sapi
38	74.6	7.6	101083	2	AC139010	AC139010 Homo sapi
39	74.6	7.6	198344	2	AC087872	AC087872 Mus muscu
40	74.6	7.6	252889	2	AC079433	AC079433 Mus muscu
41	74.4	7.5	135119	2	AC011578	AC011578 Homo sapi
42	74.2	7.5	1143	11	PM12A11G	AL684286 Penicilli
43	74.2	7.5	171574	2	AC012300	AC012300 Homo sapi
44	74	7.5	55939	2	AC091017	AC091017 Homo sapi
45	74	7.5	63082	2	AC022663	AC022663 Homo sapi

ALIGNMENTS

RESULT 1	HSMPFH1	3289 bp	DNA	linear	PRI 14-MAY-1997
LOCUS	HSMPFH1				
DEFINITION	H.sapiens MFH-1 gene.				
ACCESSION	Y08223				
VERSION	Y08223.1				
KEYWORDS	mesenchyme fork head-1 protein; MFH-1 gene.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	Miura N., Iida K., Kakinuma H., Yang X.L. and Sugiyama T.				
	Isolation of the mouse (MFH-1) and human (FXHL 14) mesenchyme fork				
	head-1 genes reveals conservation of their gene and protein				



structures  
Genomics 41 (3), 489-492 (1997)  
MEDLINE  
97312712  
PUBMED  
9169153  
REFERENCE  
2 (bases 1 to 3289)  
Miura, N.  
Direct Submission  
Submitted (18-SEP-1996) N. Miura, Akita University School of  
Medicine, Department of Biochemistry, 1-1-1 Hondo, Akita 010, JAPAN  
Location/Qualifiers  
1. 3289  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1197..2702  
/gene="MFH-1"  
1197..2702  
/gene="MFH-1"  
/codon\_start=1  
/product="Mesenchyme Fork Head-1"  
/protein\_id="CAA69400.1"  
/db\_xref="GI:1869805"  
/db\_xref="GOA:Q99958"  
/translation="MQARYSDPNALGVLPYSEONYRAAGSYGGMASPMGVSYGH  
PEYAGMGRSVAPVHHOPAPKDLVKPPYSYIALITWAIQNAPEKKITLNGIYQFI  
MDREPPYKNGKQVNSIRHNLNECFVKVPKDGKSYMTLDPDSVYKMPENG  
FLRRRRPKKQVSKERERHLEPPPAASKGAPATPLADAKAEKLVVIXSEA  
SPALVITKVELTSFESALQGSFSAATPAGSPDGSLEPHHAAPNGLPGFVENIM  
TLRTSPGGEISFGAGRAAGVVPPLALPYAAPAYGPPCAQGLEAAGAGGYOCMSR  
AMSLYTGAERFAHMCVPFALDEALSDHPSGPTSPLSALNLAAGQEGALAAATGHHQHH  
GHHHPQAPPPTPOPPQAAAAQAASQVNLNHSGLNLHPHTFAAQOQTFPNV  
REMFNSHRLGIENSLGSESQVSNASCOLPYRSTPLVRYHAAPYSYDCTKY"

ORIGIN

Query Match 100.0%; Score 986; DB 9; Length 3289;  
Best Local Similarity 100.0%; Pred. No. 4.5e-169; Indels 0; Gaps 0;  
Matches 986; Conservative 0; Mismatches 0;  
QY 1 CTGCCATTCCCAATCCAGCGCGTTGCTTTTGAATCCATTACACCTGGGCCCCCAATAA 60  
DB 212 CTGCCATTCCCAATCCAGCGCGTTGCTTTTGAATCCATTACACCTGGGCCCCCAATAA 271  
QY 61 GGAATCTAATTTATTCGTTTCATCTCATCTAATTAAGAAAATGTCACGAGNATCTGCT 120  
DB 272 GGAATCTAATTTATTCGTTTCATCTCATCTAATTAAGAAAATGTCACGAGNATCTGCT 331  
QY 121 ACTTACAAGTCTTTGGGAGAGATATTTACTCTATTAATCCATTCTATTTATATTTCA 180  
DB 332 ACTTACAAGTCTTTGGGAGAGATATTTACTCTATTAATCCATTCTATTTATATTTCA 391  
QY 181 AATTGATTTTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGGGATGGGCCCAT 240  
DB 392 AATTGATTTTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGGGATGGGCCCAT 451  
QY 241 TCACCAAAATGCATCATATAAATTAATTTAATAGATATACTTTTAAAGTTTCA 300  
DB 452 TCACCAAAATGCATCATATAAATTTAATAGATATACTTTTAAAGTTTCA 511  
QY 301 AGTGAAGAGGAGTTCGCCCGAGGCGCGGGGCTCTTAGAGCGAGGATTCCT 360  
DB 512 AGTGAAGAGGAGTTCGCCCGAGGCGCGGGGCTCTTAGAGCGAGGATTCCT 571  
QY 361 GGGCTCTGCCCCGATGGCGCGGACTCTCTAGCTGCGGGGTGATGCTCAAGT 420  
DB 572 GGGCTCTGCCCCGATGGCGCGGACTCTCTAGCTGCGGGGTGATGCTCAAGT 631  
QY 421 TCCGGAGGGGGCGTGGCGCGAGAGAAAGTAAACCTCGCTTTTACAGAAAGACTTTGA 480  
DB 632 TCCGGAGGGGGCGTGGCGCGAGAGAAAGTAAACCTCGCTTTTACAGAAAGACTTTGA 691  
481 AACTTTTCCCAATCCCTAAAGGAGCTTGGCTCTTTTCTGGGCTCAGCGGGGAGCGG 540

Db 692 AACTTTTCCCAATCCCTAAAGGAGCTTGGCTCTTTTCTGGGCTCAGCGGGGAGCGG 751  
QY 541 CTGCGAACCCCGCGCGCTGACCTCTCGGGCTCCGATTCGCTGGGGCTTGGAGAGCCTC 600  
Db 752 CTGCGAACCCCGCGCGCTGACCTCTCGGGCTCCGATTCGCTGGGGCTTGGAGAGCCTC 811  
QY 601 CTGCGAACCCCTCTCTCGCGCGCGCGGAGGCTCCACCTTGGTCCCGAGCGCGGCTCTCG 660  
Db 812 CTGCGAACCCCTCTCTCGCGCGCGCGGAGGCTCCACCTTGGTCCCGAGCGCGGCTCTCG 871  
QY 661 CTGCGGTCTCGCGCGCGCGCTCTCGCGCGCTCGCGCGCGCGGCTCTGAGCAGCGAGG 720  
Db 872 CTGCGGTCTCGCGCGCGCGCTCTCGCGCGCTCGCGCGCGCGGCTCTGAGCAGCGAGG 931  
QY 721 AGCGGGGCGCGCGCTGCTGCTGCGGGGCGCGCTCCAGATGCGGATCCGCGCGCTC 780  
Db 932 AGCGGGGCGCGCGCTGCTGCTGCGGGGCGCGCTCCAGATGCGGATCCGCGCGCTC 991  
QY 781 CGCTGAAGAGCGCGCGCTCTGCTGCGCGCGCGAGCAGCAGCGCGCACCTCGCGCGCG 840  
Db 992 CGCTGAAGAGCGCGCGCTCTGCTGCGCGCGCGAGCAGCAGCGCGCACCTCGCGCGCG 1051  
QY 841 AGCTCTCCAGAGAGCGCGCGCGCTCTGCTGCGCGCGCGCGCTCTGCTGCGCGCTCT 900  
Db 1052 AGCTCTCCAGAGAGCGCGCGCGCTCTGCTGCGCGCGCGCGCTCTGCTGCGCGCTCT 1111  
QY 901 CGCGCTCTCTCGCTCTCAGCGCGCGCGCTCTGCTGCGCGCGCGCGCTCTGCTGCGCG 960  
Db 1112 CGCGCTCTCTCGCTCTCAGCGCGCGCGCTCTGCTGCGCGCGCGCGCTCTGCTGCG 1171  
QY 961 GCGCGGAGCGCTCTCGGAAGAGCA 986  
Db 1172 GCGCGGAGCGCTCTCGGAAGAGCA 1197

RESULT 2  
AC009108/c  
LOCUS  
DEFINITION  
AC009108  
ACCESSION  
VERSION  
AC009108.10 GI:24418066  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 168656)  
AUTHORS  
DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE  
Direct Submission  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 168656)  
AUTHORS  
DOE Joint Genome Institute.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
3 (bases 1 to 168656)  
AUTHORS  
DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (29-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT  
On Oct 29, 2002 this sequence version replaced gi:13786306.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos  
National Laboratory  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.2.  
FEATURES  
source  
1. 168656  
/organism="Homo sapiens"

Query Match 100.0%; Score 12; DB 3; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
DB 112 GGAAGTAAAAA 101

RESULT 6  
US-09-134-000C-1782/c  
; Sequence 1782, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1782  
; LENGTH: 381  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-1782

Query Match 100.0%; Score 12; DB 4; Length 381;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
DB 67 GGAAGTAAAAA 56

RESULT 7  
US-09-621-976-9538  
; Sequence 9538, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9538  
; LENGTH: 535  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-9538

Query Match 100.0%; Score 12; DB 4; Length 535;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
DB 513 GGAAGTAAAAA 524

RESULT 8  
US-09-918-686-7  
; Sequence 7, Application US/09918686  
; Patent No. 6475739

Query Match 100.0%; Score 12; DB 4; Length 602;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
DB 118 GGAAGTAAAAA 129

RESULT 9  
US-09-328-352-713  
; Sequence 713, Application US/09328352  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 713  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-713

Query Match 100.0%; Score 12; DB 4; Length 678;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
DB 321 GGAAGTAAAAA 332

RESULT 10  
US-09-671-317-429  
; Sequence 429, Application US/09671317  
; Patent No. 6528260  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM  
; FILE REFERENCE: 62 US3.CIP  
; CURRENT APPLICATION NUMBER: US/09/671,317  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US 09/536,178  
; PRIOR FILING DATE: 2000-03-23  
; PRIOR APPLICATION NUMBER: PCI/IB00/00403  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: US 60/126,269  
; PRIOR FILING DATE: 1999-03-25

## RESULT 11

US-08-512-681-19/c  
; Sequence 19, Application US/08512681  
; Patent No. 5795976  
; GENERAL INFORMATION:  
; APPLICANT: Oefner, Peter A.  
; APPLICANT: Underhill, Peter A.  
; TITLE OF INVENTION: Detection of DNA Heteroduplex Molecules  
; TITLE OF INVENTION: by Denaturing High Performance Liquid Chromatography and  
; TITLE OF INVENTION: Methods for Comparative Sequencing  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/512,681  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Evans, Susan T.  
; REGISTRATION NUMBER: 38,443  
; REFERENCE/DOCKET NUMBER: 8600-0155  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: DYS234 REVERSE PRIMER  
US-08-512-681-19

Query Match 100.0%; Score 9; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9  
|||  
Db 14 ACAAATGTT 6

## RESULT 12

US-08-512-028-35  
; Sequence 15, Application US/08522028  
; Patent No. 5993813  
; GENERAL INFORMATION:  
; APPLICANT: MEZES, PETER S.  
; APPLICANT: COHEN, BRIAN B.  
; APPLICANT: RIXON, MARK W.  
; APPLICANT: ANDERSON, WH KERR  
; APPLICANT: KAPLAN, DONALD A.  
; APPLICANT: SCHOLON, JEFFREY  
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,  
; TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DUANE C ULMER  
; STREET: P.O. BOX 1967  
; CITY: MIDLAND  
; STATE: MICHIGAN  
; COUNTRY: USA

ZIP: 48641-1967  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/822,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/040,687  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ULMER, DUANE C  
; REGISTRATION NUMBER: 34,941  
; REFERENCE/DOCKET NUMBER: C-37,075C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 636-8104  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-822-028-35

Query Match 100.0%; Score 9; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9  
|||  
Db 12 ACAAATGTT 20

## RESULT 13

US-08-743-637B-240/c  
; Sequence 240, Application US/08743637B  
; Patent No. 5994066  
; GENERAL INFORMATION:  
; APPLICANT: BERGERON, Michel G.  
; APPLICANT: PICARD, Francois J.  
; APPLICANT: QUELLETTE, Marc  
; APPLICANT: ROY, Paul H.  
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA  
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND  
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED  
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...  
; NUMBER OF SEQUENCES: 273  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: QUARLES & BRADY  
; STREET: 411 EAST WISCONSIN AVENUE  
; CITY: MILWAUKEE  
; STATE: WISCONSIN  
; COUNTRY: USA  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/743,637B  
; FILING DATE: 04-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/525,840  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BAKER, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 850586.90012

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 14.7262 Seconds

(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375

Perfect score: 17

Sequence: 1 tgggaaggaataata 17

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504946 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	100.0	962	15	US-10-027-632-9718
2	17	100.0	6458	9	Sequence 9718, Ap
3	16	94.1	493	12	US-09-963-285-1
4	16	94.1	2523	12	US-10-424-599-107702
5	16	94.1	21619	10	Sequence 82796, A
6	16	94.1	21619	14	US-09-764-891-10003
7	15.4	90.6	50	15	US-10-025-428-977
8	15.4	90.6	359	12	Sequence 1138, Ap
9	15.4	90.6	386	10	US-10-424-599-64036
10	15.4	90.6	493	12	Sequence 17187, A
11	15.4	90.6	589	15	US-10-027-632-242033
12	15.4	90.6	589	15	Sequence 242034, A
13	15.4	90.6	606	12	US-10-027-632-242034
14	15.4	90.6	616	12	Sequence 141232, A
15	15.4	90.6	676	15	US-10-424-599-63676
					Sequence 148185, A

Sequence 37752, A  
Sequence 43641, A  
Sequence 256262,  
Sequence 256263,  
Sequence 256264,  
Sequence 95, Appl  
Sequence 24, Appl  
Sequence 12, Appl  
Sequence 17, Appl  
Sequence 27, Appl  
Sequence 2121, Ap  
Sequence 5, Appl  
Sequence 351, App  
Sequence 13766, A  
Sequence 68, Appl  
Sequence 1367, Ap  
Sequence 1367, Ap  
Sequence 1368, Ap  
Sequence 1369, Ap  
Sequence 1368, Ap  
Sequence 1369, Ap  
Sequence 959, App  
Sequence 1, Appl  
Sequence 6, Appl  
Sequence 9, Appl  
Sequence 1, Appl  
Sequence 8223, Ap  
Sequence 1, Appl  
Sequence 234237,

ALIGNMENTS

RESULT 1  
US-10-027-632-9718  
Sequence 9718, Application US10027632  
Publication No. US20030204075A9 ---  
GENERAL INFORMATION

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9718  
LENGTH: 962  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-9718

Query Match 100.0%; Score 17; DB 15; Length 962;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAGGAATAATA 17

```
Db 727 GTGGAAGGAATAAATA 743
|||||
RESULT 2
US-09-963-285-1
; Sequence 1, Application US/09963285
; Patent No. US20020090707A1
; GENERAL INFORMATION:
; APPLICANT: Enerback, Sven
; APPLICANT: Krock, Katarina
; APPLICANT: Rondahl, Lena
; APPLICANT: Wasserman, Wyeth
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-042001
; CURRENT APPLICATION NUMBER: US/09/963,285
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: SE 0004102-0
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/238,897
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: SE 0003435-5
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2235)...(3737)
US-09-963-285-1

Query Match 100.0%; Score 17; DB 9; Length 6458;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGAAGGAATAAATA 17
|||||
Db 359 GTGGAAGGAATAAATA 375
|||||

RESULT 3
US-10-424-599-107702
; Sequence 107702, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 107702
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(493)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68271C.1
US-10-424-599-107702

Query Match 94.1%; Score 16; DB 12; Length 493;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 727 GTGGAAGGAATAAATA 743
|||||
RESULT 4
US-10-424-599-82796
; Sequence 82796, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 82796
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45782C.1
US-10-424-599-82796

Query Match 94.1%; Score 16; DB 12; Length 2523;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGGAAGGAATAAATA 17
|||||
Db 602 GTGGAAGGAATAAATA 617
|||||

RESULT 5
US-09-764-891-10003/c
; Sequence 10003, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10003
; LENGTH: 21619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10003

Query Match 94.1%; Score 16; DB 10; Length 21619;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGAAGGAATAAATA 16
|||||
Db 5310 GTGGAAGGAATAAATA 5295
|||||

RESULT 6
US-10-205-428-977/c
; Sequence 977, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

Page 1

JOURNAL Patent: US 5635617-A 34 03-JUN-1997;

```

FEATURES
  source
    Location/Qualifiers
      1..19
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
  Query Match
    100.0%; Score 12; DB 6; Length 19;
  Best Local Similarity
    100.0%; Pred. No. 5.9e+04;
  Matches
    12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAGTAAAAA 12
Db 14 GGAAGTAAAAA 3
RESULT 2
AR147023/c
LOCUS
  Sequence 84 from patent US 6221361.
  DEFINITION
    AR147023
  ACCESSION
    AR147023.1 GI:15110826
  VERSION
    AR147023.1
  KEYWORDS
    Unknown.
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 28)
    Cochran,M.D. and Junker,D.E.
    Recombinant swinepox virus
    TITLE
    Recombinant swinepox virus
    JOURNAL
    Patent: US 6221361-A 84 24-APR-2001;
  FEATURES
    source
      Location/Qualifiers
        1..28
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match
    100.0%; Score 12; DB 6; Length 28;
  Best Local Similarity
    100.0%; Pred. No. 5.6e+04;
  Matches
    12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAGTAAAAA 12
Db 27 GGAAGTAAAAA 16
RESULT 3
BD273541/c
LOCUS
  Recombinant virus expressing foreign DNA encoding feline CD80,
  feline CD86, feline CTLA-4 or feline interferon-gamma and uses
  thereof.
  DEFINITION
    BD273541
  ACCESSION
    BD273541.1 GI:33083309
  VERSION
    JP 2002513581-A/75.
  KEYWORDS
    unidentified
  SOURCE
    unidentified
  ORGANISM
    unclassified.
  REFERENCE
    1 (bases 1 to 28)
    Winslow,B.J. and Cochran,M.D.
    Recombinant virus expressing foreign DNA encoding feline CD80,
    feline CD86, feline CTLA-4 or feline interferon-gamma and uses
    Patent: JP 2002513581-A 75 14-MAY-2002;
  JOURNAL
    SCHERING-PLOUGH LTD
  COMMENT
    OS Unidentified
    PN JP 2002513581-A/75
    PD 14-MAY-2002
    PF 30-APR-1999 JP 2000547248
    PR 01-MAY-1998 US 09/071711
    PI BARBARA J WINSLOW,MARK D COCHRAN
    PC C12N15/09,A61K39/12,A61K39/15,A61K39/215,A61K39/23,
    PC A61K39/245,
    PC A61K48/00,A61P43/00,C12N7/00//C07K14/705,C12N15/00 CC FIV
    PPR downstream primer
    PH* Key
    Location/Qualifiers
      1..28
      /organism="unidentified"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32644"
ORIGIN
  Query Match
    100.0%; Score 12; DB 6; Length 28;
  Best Local Similarity
    100.0%; Pred. No. 5.6e+04;
  Matches
    12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAGTAAAAA 12
Db 27 GGAAGTAAAAA 16
RESULT 4
AR267926/c
LOCUS
  Sequence 265 from patent US 6497882.
  DEFINITION
    AR267926
  ACCESSION
    AR267926
  VERSION
    AR267926.1 GI:29698051
  KEYWORDS
    Unknown.
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 28)
    Cochran,M.D. and Junker,D.E.
    Recombinant swinepox virus
    TITLE
    Recombinant swinepox virus
    JOURNAL
    Patent: US 6497882-A 265 24-DEC-2002;
  FEATURES
    source
      Location/Qualifiers
        1..28
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
  Query Match
    100.0%; Score 12; DB 6; Length 28;
  Best Local Similarity
    100.0%; Pred. No. 5.6e+04;
  Matches
    12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAGTAAAAA 12
Db 27 GGAAGTAAAAA 16
RESULT 5
AX093873
LOCUS
  Sequence 2 from Patent WO0118039.
  DEFINITION
    AX093873
  ACCESSION
    AX093873
  VERSION
    AX093873.1 GI:13510091
  KEYWORDS
    Homo sapiens (human)
  SOURCE
    Homo sapiens
  ORGANISM
    Homo sapiens
  REFERENCE
    Pitha-Roe,P.M., Yeow,W.S. and Au,W.C.
    Increased production of interferon- $\gamma$ (a)
    Patent: WO 0118039-A 2 15-MAR-2001;
  JOURNAL
    GLAXO GROUP LIMITED (GB)
  FEATURES
    source
      Location/Qualifiers
        1..58
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
  Query Match
    100.0%; Score 12; DB 6; Length 58;
  Best Local Similarity
    100.0%; Pred. No. 5e+04;
  Matches
    12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAGTAAAAA 12
Db 27 GGAAGTAAAAA 16

```

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 12 GGAAGTAAAAA 23

RESULT 6  
 AF227460/c 150 bp DNA linear PLN 13-JUN-2000  
 LOCUS Bromus tectorum clone 21 microsatellite sequence.  
 DEFINITION  
 ACCESSION AF227460  
 VERSION AF227460.1 GI:8489262

KEYWORDS Bromus tectorum  
 SOURCE Bromus tectorum  
 ORGANISM Bromus tectorum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Bromaceae; Bromus.

REFERENCE 1 (bases 1 to 150)  
 AUTHORS Paulsen, A., Meyer, S.E., Coleman, C.E. and Fairbanks, D.J.  
 TITLE Microsatellite markers for Bromus tectorum (cheatgrass)  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 150)  
 AUTHORS Paulsen, A., Meyer, S.E., Coleman, C.E. and Fairbanks, D.J.  
 TITLE Direct Submission

JOURNAL Submitted (21-JAN-2000) Botany and Range Science, Brigham Young University, 401 WIDB, Provo, UT 84602, USA

FEATURES  
 source 1..150  
 /organism="Bromus tectorum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:29867"  
 /clone="21"  
 repeat\_region 1..150  
 /notes="microsatellite"  
 /rpt\_type=tandem

## ORIGIN

Query Match 100.0%; Score 12; DB 8; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+04; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 58 GGAAGTAAAAA 47

RESULT 7  
 AX305717 151 bp DNA linear PAT 11-DEC-2001  
 LOCUS Sequence 468 from Patent WO0188188.  
 DEFINITION  
 ACCESSION AX305717

VERSION AX305717.1 GI:17645138  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
 Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.  
 TITLE Method for examining ischemic conditions

JOURNAL Patent: WO 0188188-A 468 22-NOV-2001;  
 School Juridical Person Nihon University (JP)

FEATURES  
 source 1..151  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+04; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 102 GGAAGTAAAAA 113

RESULT 8  
 BX322143/c 191 bp DNA linear STS 10-JUN-2003  
 LOCUS Arabidopsis thaliana transposon insertion STS SM\_3.35925, sequence tagged site.

DEFINITION  
 ACCESSION BX322143  
 VERSION BX322143.1 GI:29420274  
 KEYWORDS STS; sequence tagged site.  
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1  
 AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.

JOURNAL Unpublished  
 2 (bases 1 to 191)  
 Clarke, J.H.

REFERENCE Direct Submission  
 TITLE Submitted (31-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

JOURNAL AT denotes an activation tag dissociation transposon within a single line. ET an enhancer trap dissociation transposon. GT a gene trap dissociation transposon. MT a mis-expression enhancer trap dissociation transposon. SM a defective suppressor mutator transposon. \_3 denotes a sequence derived from the 3' end of the transposon. \_5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project  
 On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: N12836.

## FEATURES

source 1..191  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /variety="Columbia-0 NASC stock code N1092"  
 /db\_xref="taxon:3702"  
 /clone="AC006841"  
 /note="Derived from superpool 25.06 NASC code N41149"  
 1..191  
 /standard\_name="SM\_3.35925"

## ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 89 GGAAGTAAAAA 78

## RESULT 9

G55264/c 206 bp DNA linear STS 30-MAR-2000

LOCUS SHGC-100532 Human Homo sapiens STS genomic, sequence tagged site.

DEFINITION

ACCESSION G55264

VERSION G55264.1 GI:6120583

KEYWORDS STS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 206)

AUTHORS Olivier, M. and Cox, D.R.

TITLE Unpublished, Olivier, M., Cox, D.R. (2000)

JOURNAL Unpublished (2000)



## COMMENT

Contact: Michael Olivier, David R. Cox  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
 Tel: (650) 320-5800  
 Fax: (650) 320-5801  
 Email: olivier@hgc.stanford.edu  
 Primer A: TCAGTTAACTTTCGAAACCCATT  
 Primer B: CTAGTGTTCCTTCCTTCCTCCT  
 SFS size: 127  
 PCR profile:

Initial incubation: 95 degrees C for 10 minutes  
 Denaturation: 94 degrees C for 30 seconds  
 Annealing: 60 degrees C for 30 seconds  
 Polymerization: 72 degrees C for 23 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9700  
 Protocol:  
 Template: 25 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 AmpliTaq Gold Polymerase: 0.07 units/ul  
 Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 8.3

BAC ends sequenced at TIGR from the RC111 BAC library. Designed and developed at the Stanford Human Genome Center.

FEATURES  
source

Location/Qualifiers  
 1..206  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map\_d="14"  
 /clone\_lib="Human"  
 19..145  
 primer\_bind  
 19..41  
 complement(123..145)

## ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 196 GGAAGTAAAAA 185

## RESULT 10

AX909385  
 LOCUS AX909385 25248 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 25248 from Patent EP1033401.  
 ACCESSION AX909385  
 VERSION AX909385.1 GI:40065465  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1  
 Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.  
 Expressed sequence tags and encoded human proteins  
 Patent: EP 1033401-A 25248 06-SEP-2000;  
 Genset (FR)

FEATURES  
source

Location/Qualifiers  
 1..207  
 /organism="Homo sapiens"

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 79 GGAAGTAAAAA 90

## RESULT 11

BD044918  
 LOCUS BD044918 207 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Sequence tag and encoded human protein.  
 ACCESSION BD044918  
 VERSION BD044918.1 GI:22586660  
 KEYWORDS JP 2001269182-A/21164.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 207)  
 Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
 Sequence tag and encoded human protein  
 Patent: JP 2001269182-A 21164 02-OCT-2001;  
 Genset

## COMMENT

OS Homo sapiens (human)  
 PN JP 2001269182-A/21164  
 PD 02-OCT-2001  
 PF 24-FEB-2000 JP 2000118773  
 PR 26-FEB-1999 US 60/122487  
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES  
 PI JORDAN  
 PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC  
 C12N5/10.  
 PC C12P21/02, C12P21/08, C12Q1/68 //G06F17/30, C12N15/00, C12N5/00, PC  
 G06F15/40  
 CC  
 FH Key Location/Qualifiers  
 1..207  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

FEATURES  
source

Location/Qualifiers  
 1..207  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 79 GGAAGTAAAAA 90

## RESULT 12

BNAS78474  
 LOCUS BNAS78474 209 bp DNA linear PLN 25-OCT-2003  
 DEFINITION Brassica napus chloroplast partial psbc gene for PSII 43 kDa  
 protein, IGS and partial trna-Ser (UGA) gene.  
 ACCESSION AJ578474  
 VERSION AJ578474.1 GI:37992039  
 KEYWORDS IGS; intergenic spacer; psbc gene; PSII 43 kDa protein; transfer  
 RNA-Ser (UGA); trna-Ser (UGA) gene.  
 SOURCE Chloroplast Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

## REFERENCE

1

AUTHORS Zubko,M.K., Zubko,E.I., van Zuijlen,K., Meyer,P. and Day,A.  
 TITLE Stable plasmid transformation of Petunia plastids  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 209)  
 AUTHORS Day,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-JUL-2003) Day A., School of Biological Sciences,  
 University of Manchester, 3.614 Stopford Building, Oxford Road, M13  
 9PT, UNITED KINGDOM

FEATURES  
 source  
 1..209  
 /organism="Brassica napus"  
 /organella="plastid;chloroplast"  
 /mol\_type="genomic DNA"  
 /cultivar="nars"  
 /db\_xref="taxon:3708"  
 1..18  
 /genes="pabC"  
 <1..18  
 /genes="pabC"  
 /function="photosynthetic protein"  
 /codon\_start=1  
 /transl\_table=11  
 /product="PSII 43 kDa protein"  
 /protein\_id="CA18106.1"  
 /db\_xref="GI:37992040"  
 /translations="MTPLN"  
 19..193  
 /notes="intergenic spacer, IGS"  
 complement(194..209)  
 /genes="trna-Ser (UGA)"  
 complement(194..>209)  
 /genes="trna-Ser (UGA)"  
 /product="trna-Ser"  
 /notes="codon recognized: UCA"

misc\_feature  
 19..193  
 /notes="intergenic spacer, IGS"  
 complement(194..209)  
 /genes="trna-Ser (UGA)"  
 complement(194..>209)  
 /genes="trna-Ser (UGA)"  
 /product="trna-Ser"  
 /notes="codon recognized: UCA"

gene  
 194..209  
 /genes="trna-Ser (UGA)"  
 complement(194..>209)  
 /genes="trna-Ser (UGA)"  
 /product="trna-Ser"  
 /notes="codon recognized: UCA"

CDS  
 194..209  
 /genes="trna-Ser (UGA)"  
 complement(194..>209)  
 /genes="trna-Ser (UGA)"  
 /product="trna-Ser"  
 /notes="codon recognized: UCA"

ORIGIN  
 Query Match 100.0%; Score 12; DB 8; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 Db 35 GGAAAGTAAAAA 46

RESULT 13  
 HSPE10B10/c  
 LOCUS HSPE10B10 210 bp DNA linear STS 21-MAY-1999  
 DEFINITION H.sapiens flow-sorted chromosome 1 HindIII fragment, SCIP10B10,  
 sequence tagged site.  
 ACCESSION AL009426  
 VERSION AL009426.1 GI:2664589  
 KEYWORDS STS; single read.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 210)  
 AUTHORS Gregory,S., Kettleborough,R., Langford,C., Ross,M.T. and Hunt,S.E.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-DEC-1997) E-mail contact: humquery@sanger.ac.uk  
 COMMENT Vector: pBS1ISK+ (Primer A : CTCATCTCACTCTGTCCTCA; Primer B :  
 Marker scg29044 (Primer A : CTCATCTCACTCTGTCCTCA; Primer B :  
 TCATAAGCCCACTAGTGTGA; amplicon size : 143 bp) was mapped to  
 chromosome 1 using Radiation Hybrid  
 panel Genebridge 4 (GB4).  
 Location/Qualifiers  
 source  
 1..210  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"

FEATURES  
 source  
 1..210  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"

ORIGIN  
 Query Match 100.0%; Score 12; DB 8; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 Db 35 GGAAAGTAAAAA 46

RESULT 14  
 AF184852/c  
 LOCUS AF184852 229 bp DNA linear PLN 24-APR-2001  
 DEFINITION Arabidopsis thaliana inhibitor tagged site ITS85 genomic sequence.  
 ACCESSION AF184852  
 VERSION AF184852.1 GI:6272782  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 REFERENCE 1 (bases 1 to 229)  
 AUTHORS Speulman,E., Metz,P.L., van Arkel,G., te Lintel Hekkert,B.,  
 Stiekema,W.J. and Pereira,A.  
 TITLE A two-component enhancer-inhibitor transposon mutagenesis system  
 for functional analysis of the Arabidopsis genome  
 JOURNAL Plant Cell 11 (10), 1853-1866 (1999)  
 MEDLINE 99452723  
 PUBMED 10521517  
 REFERENCE 2 (bases 1 to 229)  
 AUTHORS Speulman,E., Metz,P.L., van Arkel,G., te Lintel Hekkert,B.,  
 Stiekema,W.J. and Pereira,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-SEP-1999) Molecular Biology, CPRO-DLO, P.O. Box 16,  
 Wageningen 6700AA, The Netherlands  
 Location/Qualifiers  
 source  
 1..229  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 /clone="ITS85"

ORIGIN  
 Query Match 100.0%; Score 12; DB 8; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 Db 184 GGAAAGTAAAAA 173

RESULT 15  
 AU046405  
 LOCUS AU046405 247 bp DNA linear STS 22-APR-1999  
 DEFINITION Rattus norvegicus, OTSUKA clone, 01478, microsatellite sequence,  
 sequence tagged site.  
 ACCESSION AU046405  
 VERSION AU046405.1 GI:4631041  
 KEYWORDS STS.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (sites)

ORIGIN  
 Query Match 100.0%; Score 12; DB 11; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 Db 21 GGAAAGTAAAAA 10

RESULT 14  
 AF184852/c  
 LOCUS AF184852 229 bp DNA linear PLN 24-APR-2001  
 DEFINITION Arabidopsis thaliana inhibitor tagged site ITS85 genomic sequence.  
 ACCESSION AF184852  
 VERSION AF184852.1 GI:6272782  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 REFERENCE 1 (bases 1 to 229)  
 AUTHORS Speulman,E., Metz,P.L., van Arkel,G., te Lintel Hekkert,B.,  
 Stiekema,W.J. and Pereira,A.  
 TITLE A two-component enhancer-inhibitor transposon mutagenesis system  
 for functional analysis of the Arabidopsis genome  
 JOURNAL Plant Cell 11 (10), 1853-1866 (1999)  
 MEDLINE 99452723  
 PUBMED 10521517  
 REFERENCE 2 (bases 1 to 229)  
 AUTHORS Speulman,E., Metz,P.L., van Arkel,G., te Lintel Hekkert,B.,  
 Stiekema,W.J. and Pereira,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-SEP-1999) Molecular Biology, CPRO-DLO, P.O. Box 16,  
 Wageningen 6700AA, The Netherlands  
 Location/Qualifiers  
 source  
 1..229  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 /clone="ITS85"

ORIGIN  
 Query Match 100.0%; Score 12; DB 8; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 Db 184 GGAAAGTAAAAA 173

RESULT 15  
 AU046405  
 LOCUS AU046405 247 bp DNA linear STS 22-APR-1999  
 DEFINITION Rattus norvegicus, OTSUKA clone, 01478, microsatellite sequence,  
 sequence tagged site.  
 ACCESSION AU046405  
 VERSION AU046405.1 GI:4631041  
 KEYWORDS STS.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (sites)

AUTHORS Watanabe,T.K., Hishigaki,H., Okuno,S., Mizoguchi,A., Oga,K.,  
Tsuiji,A., Ono,T., Yasasaki,Y., Kanemoto,N., Takahashi,B., Irie,Y.,  
Nakamura,Y., Takagi,Y. and Tanigami,A.  
TITLE The large-scale mapping of rat microsatellite markers  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 247)  
AUTHORS Watanabe,T.K.  
TITLE Direct Submission  
JOURNAL Submitted (11-DEC-1998) Takeshi K Watanabe, Otsuka GEN Research  
Institute, Otsuka Pharmaceutical Co., Ltd, 463-10, Kagasuno,  
Kawauchi-cho, Tokushima, Tokushima 771-0192, Japan  
(E-mail:watanabe@otsuka.gr.jp, Tel:+81-886-65-2888,  
Fax:+81-886-37-1035)  
FEATURES  
source Location/Qualifiers  
1..247  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="Brown Norway"  
/db\_xref="taxon:10116"  
/clone="01478"  
/cell\_type="hepatocyte"  
/tissue\_type="liver"  
/notes="01478F=5'-TCACATCCACCATCTCAAT-3',  
01478R=5'-ACAACTATTACTGCCGAGCC-3'"

## ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 247;  
Best Local Similarity 100.0%; Pred.No. 4e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12  
| | | | | | | | | | | | | | | |  
Db 98 GGAAAGTAAAAA 109

Search completed: March 25, 2004, 12:49:42  
Job time : 85.7624 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:55:14 ; Search time 2.36035 Seconds  
(without alignments)  
2821.370 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1692\_1703

Perfect score: 12

Sequence: 1 ggaagtaaaaa 12

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	12	100.0	19	1	US-08-233-788A-34
C 2	12	100.0	28	3	US-08-686-968C-84
C 3	12	100.0	28	4	US-08-472-679H-265
C 4	12	100.0	28	4	US-09-303-040-80
C 5	12	100.0	361	3	US-09-385-982-26
C 6	12	100.0	381	4	US-09-134-000C-1782
C 7	12	100.0	535	4	US-09-621-976-9538
C 8	12	100.0	602	4	US-09-918-686-7
C 9	12	100.0	678	4	US-09-328-352-713
C 10	12	100.0	710	4	US-09-671-317-429
C 11	12	100.0	732	4	US-09-134-000C-701
C 12	12	100.0	966	4	US-09-023-655-875
C 13	12	100.0	1001	4	US-09-641-638-524
C 14	12	100.0	1001	4	US-09-641-638-525
C 15	12	100.0	1131	4	US-09-620-312D-203
C 16	12	100.0	1245	1	US-08-599-171A-17
C 17	12	100.0	1245	2	US-08-646-590B-17
C 18	12	100.0	1245	3	US-09-069-226-17
C 19	12	100.0	1245	3	US-09-412-184-17
C 20	12	100.0	1332	3	US-09-457-046B-49
C 21	12	100.0	1401	1	US-08-785-066-1
C 22	12	100.0	1401	3	US-09-007-355-1
C 23	12	100.0	1401	3	US-08-913-489-1
C 24	12	100.0	1483	1	US-08-270-076A-2
C 25	12	100.0	1801	4	US-09-620-312D-404
C 26	12	100.0	1838	5	PCT-US93-06251-85
C 27	12	100.0	2193	4	US-09-011-143-5

28	12	100.0	2193	4	US-09-302-495-5	Sequence 5, Appli
29	12	100.0	2193	4	US-10-079-616-5	Sequence 5, Appli
30	12	100.0	2370	4	US-09-873-404-1	Sequence 1, Appli
C 31	12	100.0	2497	6	5185259-2	Patent No. 5185259
32	12	100.0	3093	4	US-08-811-682-14	Sequence 14, Appli
C 33	12	100.0	3499	3	US-08-857-076-43	Sequence 43, Appli
C 34	12	100.0	3841	4	US-08-811-682-13	Sequence 13, Appli
35	12	100.0	4084	4	US-09-221-017B-632	Sequence 632, App
36	12	100.0	4406	4	US-08-233-788A-39	Sequence 39, Appli
C 37	12	100.0	6876	4	US-08-956-17E-186	Sequence 186, App
38	12	100.0	9468	1	US-08-325-547-10	Sequence 10, Appli
39	12	100.0	9471	4	US-08-811-682-1	Sequence 1, Appli
40	12	100.0	9472	1	US-08-325-547-9	Sequence 9, Appli
41	12	100.0	9751	3	US-09-238-303-7	Sequence 7, Appli
42	12	100.0	9751	4	US-09-946-239-7	Sequence 7, Appli
43	12	100.0	13473	5	PCT-US96-03916-1	Sequence 1, Appli
44	12	100.0	18554	4	US-09-811-825A-3	Sequence 3, Appli
45	12	100.0	18912	5	PCT-US96-03916-59	Sequence 59, Appli

ALIGNMENTS

RESULT 1  
US-08-233-788A-34/c  
; Sequence 34, Application US/08233788A  
; Patent No. 5635617  
; GENERAL INFORMATION:  
; APPLICANT: Doran, James L.  
; APPLICANT: Kay, William W.  
; APPLICANT: Collinson, Karen S.  
; APPLICANT: Cloutier, Sharon C.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION  
; TITLE OF INVENTION: OF SALMONELLA  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/233,788A  
; FILING DATE: 26-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, Joshua  
; REGISTRATION NUMBER: 35,570  
; REFERENCE/DOCKET NUMBER: 920043.403C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836 SEDANBERRY  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Other nucleic acid  
; US-08-233-788A-34

Query Match 100.0%; Score 12; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12

```
Db      |||||
        14 GGAAGTAAAAA 3

RESULT 2
US-08-686-968C-84/c
; Sequence 84, Application US/08686968C
; Patent No. 6221361
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Jucker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; FILE REFERENCE: 39119-H/JML
; CURRENT APPLICATION NUMBER: US/08/686,968C
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 84
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
; US-08-686-968C-84

Query Match      100.0%; Score 12; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAAGTAAAAA 12
        |||||
Db      27 GGAAGTAAAAA 16

RESULT 3
US-08-472-679H-265/c
; Sequence 265, Application US/08472679H
; Patent No. 6497882
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Jucker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 267
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pamela G. Salkeld
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,679H
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Salkeld, Pamela G.
; REGISTRATION NUMBER: 38,607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 298-2135
; TELEFAX: (908) 298-5388
; INFORMATION FOR SEQ ID NO: 265:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (Genomic)
; HYPOTHETICAL: NO
```

```
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 265:
US-08-472-679H-265

Query Match      100.0%; Score 12; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAAGTAAAAA 12
        |||||
Db      27 GGAAGTAAAAA 16

RESULT 4
US-09-303-040-80/c
; Sequence 80, Application US/09303040
; Patent No. 6555671
; GENERAL INFORMATION:
; APPLICANT: Winslow, Barbara J.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: Recombinant Virus Expressing Foreign DNA Encoding
; TITLE OF INVENTION: Feline CD80, Feline CD86, Feline CD28, Feline CD134 or
; TITLE OF INVENTION: Feline Interferon-gamma And Uses Thereof
; FILE REFERENCE: 54957-B
; CURRENT APPLICATION NUMBER: US/09/303,040
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,870
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 80
; LENGTH: 28
; TYPE: DNA
; ORGANISM: FIV PPR downstream primer
; US-09-303-040-80

Query Match      100.0%; Score 12; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAAGTAAAAA 12
        |||||
Db      27 GGAAGTAAAAA 16

RESULT 5
US-09-385-982-26/c
; Sequence 26, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCNDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(361)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-26
```

```
Query Match      100.0%; Score 12; DB 3; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
   |||||
Db 112 GGAAGTAAAAA 101

RESULT 6
US-09-134-000C-1782/c
; Sequence 1782, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1782
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1782

Query Match      100.0%; Score 12; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
   |||||
Db 67 GGAAGTAAAAA 56

RESULT 7
US-09-621-976-9538
; Sequence 9538, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9538
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9538

Query Match      100.0%; Score 12; DB 4; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
   |||||
Db 513 GGAAGTAAAAA 524

RESULT 8
US-09-918-686-7
; Sequence 7, Application US/09918686
; Patent No. 6475739
```

```
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Prohl, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; FILE REFERENCE: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-686-7
```

```
Query Match      100.0%; Score 12; DB 4; Length 602;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGTAAAAA 12
   |||||
Db 118 GGAAGTAAAAA 129
```

```
RESULT 9
US-09-328-352-713
; Sequence 713, Application US/09328352
; Patent No. 6564998
```

```
; GENERAL INFORMATION:
; APPLICANT: Selby L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 713
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-713
```

```
Query Match      100.0%; Score 12; DB 4; Length 678;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GGAAGTAAAAA 12
   |||||
Db 321 GGAAGTAAAAA 332
```

```
RESULT 10
US-09-671-317-429
; Sequence 429, Application US/09671317
; Patent No. 6528260
```

```
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/671,317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
```

```
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 429
; LENGTH: 710
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 364
; OTHER INFORMATION: 12-622-325 : polymorphic base C or T
; NAME/KEY: misc binding
; LOCATION: 344..363
; OTHER INFORMATION: 12-622-325.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 365..384
; OTHER INFORMATION: 12-622-325.misl, potential complement
; NAME/KEY: primer_bind
; LOCATION: 40..59
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 551..569
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 352..376
; OTHER INFORMATION: 12-622-325 potential probe
; NAME/KEY: misc feature
; LOCATION: 333,685,688
; OTHER INFORMATION: n=a, g, c or t
; US-09-671-317-429

Query Match 100.0%; Score 12; DB 4; Length 710;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
Db 668 GGAAGTAAAAA 679

RESULT 11
US-09-134-000C-701/C
; Sequence 701, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 701
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-701

Query Match 100.0%; Score 12; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
Db 576 GGAAGTAAAAA 565

RESULT 12
US-09-023-655-875
; Sequence 875, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 875:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1088447
; US-09-023-655-875

Query Match 100.0%; Score 12; DB 4; Length 966;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
Db 249 GGAAGTAAAAA 260

RESULT 13
US-09-641-638-524
; Sequence 524, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
```

```
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 524
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-298-122 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-298-122.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-298-122.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 384..401
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 806..825
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-298-122 potential probe
; US-09-641-638-524

Query Match 100.0%; Score 12; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
Db 575 GGAAAGTAAAAA 586

RESULT 14
US-09-641-638-525
; Sequence 525, Application US/09641638
; Patent No. 6432848
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 525
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-298-158 : polymorphic base A or G
```

```
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-298-158.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-298-158.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 348..365
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 770..789
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-298-158 potential probe
; US-09-641-638-525

Query Match 100.0%; Score 12; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
Db 539 GGAAAGTAAAAA 550

RESULT 15
US-09-620-312D-203
; Sequence 203, Application US/09620312D
; Patent No. 6569862
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 203
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)..(1082)
; US-09-620-312D-203

Query Match 100.0%; Score 12; DB 4; Length 1131;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
```



Db 308 GGAAGTAAAA 319  
|||||

Search completed: March 25, 2004, 15:34:19  
Job time : 4.36035 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 08:39:03 ; Search time 6718.14 Seconds  
(without alignments)  
6361.316 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1250\_2235

Perfect score: 986

Sequence: 1 ctgccattccattccagcgc.....gagccgtctcgaagcagca 986

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pin.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	986	100.0	3289	9	HSMFH1	Y08223 H.sapiens M
c	298.6	30.3	178416	10	AC009108	AC009108 Homo sapi
3	298.6	30.3	178416	10	AC0127554	AC0127554 Mus muscu
4	290	29.4	6021	10	MMFHEAD1	Y08222 M.musculus
5	88.4	9.0	159805	2	BX546449	BX546449 Danio rer
c	83	8.4	869	11	PM2A12B	AL684671 Penicilli
c	80.6	8.2	172457	2	BX664728	BX664728 Danio rer
c	79.8	8.1	2000	6	AX655393	AX655393 Sequence
c	79.8	8.1	51753	2	AC124247	AC124247 Homo sapi
c	79.8	8.1	75144	2	AC027483	AC027483 Homo sapi
c	79.6	8.1	885	11	PM7F8G	AL685189 Penicilli
c	79.6	8.1	11305	8	QSA535061	AL685189 Penicilli
c	79.2	8.0	85434	2	AC066610	AC066610 Homo sapi
c	79	8.0	224777	2	AC138109	AC138109 Mus muscu
c	78.8	8.0	93621	2	AC021596	AC021596 Homo sapi
c	78.6	8.0	177948	2	AC146089	AC146089 Pan trogl
c	78	7.9	300695	2	AC079431	AC079431 Mus muscu
c	77.6	7.9	192042	2	AC063969	AC063969 Mus muscu
c	77.4	7.8	1279	11	PM2H12G	AL684840 Penicilli
c	77.4	7.8	181850	2	BX276102	BX276102 Danio rer
c	77	7.8	219952	2	AC084804	AC084804 Mus muscu
c	77	7.8	63540	2	AC110231	AC110231 Mus muscu
c	76.8	7.8	6931	2	AC067888	AC067888 Homo sapi
c	76.2	7.7	205691	2	AC105113	AC105113 Homo sapi
c	76.2	7.7	205691	2	AC087227	AC087227 Mus muscu
c	76	7.7	207420	2	AC078884	AC078884 Mus muscu
c	75.6	7.7	69570	2	AC067888	AC067888 Homo sapi
c	75.2	7.6	81767	2	AC021929	AC021929 Homo sapi
c	75.2	7.6	153657	2	BX469901	BX469901 Danio rer
c	75.2	7.6	216645	2	AC119156	AC119156 Mus muscu
c	75	7.6	135552	2	AC019251	AC019251 Homo sapi
c	75	7.6	190822	2	AC051613	AC051613 Mus muscu
c	75	7.6	216332	2	AC145342	AC145342 Pan trogl
c	74.8	7.6	78220	2	AC023212	AC023212 Homo sapi
c	74.8	7.6	138709	2	BX649540	BX649540 Danio rer
c	74.8	7.6	217412	2	AC024400	AC024400 Homo sapi
c	74.6	7.6	52884	2	AC068018	AC068018 Homo sapi
c	74.6	7.6	101083	2	AC139010	AC139010 Homo sapi
c	74.6	7.6	198344	2	AC087872	AC087872 Mus muscu
c	74.6	7.6	253689	2	AC079433	AC079433 Mus muscu
c	74.4	7.5	135119	2	AC011578	AC011578 Homo sapi
c	74.2	7.5	1143	11	PM12A11G	AL684286 Penicilli
c	74.2	7.5	171574	2	AC012300	AC012300 Homo sapi
c	74	7.5	55939	2	AC091017	AC091017 Homo sapi
c	74	7.5	63082	2	AC022663	AC022663 Homo sapi

ALIGNMENTS

RESULT 1

HSMFH1

LOCUS

DEFINITION H.sapiens MFH-1 gene.

ACCESSION Y08223

VERSION Y08223.1

KEYWORDS mesenchyme fork head-1 protein; MFH-1 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE Miura,N., Iida,K., Kakinuma,H., Yang,X.L. and Sugiyama,T.

AUTHORS Isolation of the mouse (MFH-1) and human (pKHL-14) mesenchyme fork

TITLE head-1 genes reveals conservation of their gene and protein

3289 bp DNA linear PRI 14-MAY-1997

JOURNAL  
MEDLINE  
PUBMED  
9169153  
REFERENCE  
2 (bases 1 to 3289)  
AUTHORS  
Miura,N.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (18-SEP-1996) N. Miura, Akita University School of  
Medicine, Department of Biochemistry, 1-1-1 Hondo, Akita 010, JAPAN  
FEATURES  
Location/Qualifiers  
1..3289  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1197..2702  
/gene="MFH-1"  
1197..2702  
/gene="MFH-1"  
/codon\_start=1  
/product="Mesenchyme Fork Head-1"  
/protein\_id="CAA69400.1"  
/db\_xref="GI:1869805"  
/db\_xref="GOA:Q99958"  
/db\_xref="SWISS-PROT:Q99958"  
/translation="MOARVSDPNALGVVLYSEQNYVRAAGSYCGMASPMGVYSGH  
PEQYAGMRGSAVYHHQPAAPKDLVFPYVYALITMAIQNAPEKKITLNGIYQFI  
MDRFPYREKQWONSHIHNLISLNECFVKVPRDDKKGKGSYMTLDPDSYNMFENG  
FLRRRRPKKQVSKERAEHLKBPAPPAASKGAPATPHLADAPKEAEKKVYKSEA  
SPALVITKVTLSPEALQSGPSRAASTPAGSDGSLPEHHAAPNGLPGFSVENIM  
TLRTSPGSELSPGAGRLVYVPLALPYAAPPAAYQPCNAGLEAAGAGYQCSMR  
ANSLYGAERPAHMCVFPALDEALSDHSGPTSPUSALNLAAGGALAAATGHHQHH  
GHHQPAAPPAPQPPQPPQAAAAAQAASWYLNHSGDLNHLPGHTFAAQOQTPNV  
REMFNSHRLGIENSLGESQVSNASCOLPYRSTPPLRYHAAPYSYDCTKY"  
ORIGIN  
Query Match 100.0%; Score 986; DB 9; Length 3289;  
Best Local Similarity 100.0%; Pred. No. 4.5e-169; Indels 0; Gaps 0;  
Matches 986; Conservative 0; Mismatches 0;  
QY 1 CTGCCATTCATCCAGCGCTTGGCTTTGATCCATTCACCTGGGCCCCCAATTA 60  
DB 212 CTGCCATTCATCCAGCGCTTGGCTTTGATCCATTCACCTGGGCCCCCAATTA 271  
QY 61 GGAATCTAATTCGCTTCATCTCACTCAATTAAGAAATATGCCCAGATCTTCT 120  
DB 272 GGAATCTAATTCGCTTCATCTCACTCAATTAAGAAATATGCCCAGATCTTCT 331  
QY 121 ACTTCAAGCTCTTGGGAGAGATATTTACTCTATTAATCCATCTATTTATATTTCA 180  
DB 332 ACTTCAAGCTCTTGGGAGAGATATTTACTCTATTAATCCATCTATTTATATTTCA 391  
QY 181 AATTGATTTTAAACAGAGAAAGTGCTATCTTTTGGTGGGATGCGGCCCAT 240  
DB 392 AATTGATTTTAAACAGAGAAAGTGCTATCTTTTGGTGGGATGCGGCCCAT 451  
QY 241 TCACCAAAATGTGATCATAAAATAAATTTAATAAGATATTAATTTAAAGTTTCA 300  
DB 452 TCACCAAAATGTGATCATAAAATAAATTTAATAAGATATTAATTTAAAGTTTCA 511  
QY 301 AATGAAGCGAGTCCCGCGAGCGCGGGCGGGGCTTTAGACCGAGATTCCT 360  
DB 512 AATGAAGCGAGTCCCGCGAGCGCGGGCGGGGCTTTAGACCGAGATTCCT 571  
QY 361 GCGCTCTCCGCCCATGCGCGCGGACCTCTCTCAGCTGCGGGTGTGCTCAAGT 420  
DB 572 GCGCTCTCCGCCCATGCGCGCGGACCTCTCTCAGCTGCGGGTGTGCTCAAGT 631  
QY 421 TCCGGAGGGGCGTGGGCCGAGGAAGTAAATACTGCTTTACCAAGAGACTTTTGA 480  
DB 632 TCCGGAGGGGCGTGGGCCGAGGAAGTAAATACTGCTTTACCAAGAGACTTTTGA 691  
QY 481 AACTTTTCCCAATCCCTAAAGGGACTTGGCTCTTTTCTGGCTCAGCGGGGAGCGG 540

Db 592 AACTTTTCCCAATCCCTAAAGGGACTTGGCTCTTTTCTGGCTCAGCGGGGAGCGG 751  
QY 541 CTCCGAGCCCCGGCGCGCTGACCCCTCGGGGCTGCGGATTCGCTCGGGGCTTTGGAGAGCTC 600  
Db 752 CTCCGAGCCCCGGCGCGCTGACCCCTCGGGGCTGCGGATTCGCTCGGGGCTTTGGAGAGCTC 811  
QY 601 CTCCGAGCCCCCTCTCGCGCGGGCGAGGGTCCACTTGGTCCCGAGCGCGGGGCTTCCG 660  
Db 812 CTCCGAGCCCCCTCTCGCGCGGGCGAGGGTCCACTTGGTCCCGAGCGCGGGGCTTCCG 871  
QY 661 CTCCGAGCCCCGGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTTCTGGAGCGAGG 720  
Db 872 CTCCGAGCCCCGGCGCGCTGCGCGCGCTGCGCGCGCTTCTGGAGCGAGG 931  
QY 721 AGCGGGGCGCGCTGCGCTTGGCGGGGCGCGCTTCCAGGATGCGATCCCGCGGCT 780  
Db 932 AGCGGGGCGCGCTGCGCTTGGCGGGGCGCGCTTCCAGGATGCGATCCCGCGGCT 991  
QY 781 CGCTGAAAGGCGCGCGCTGCTCGGCGCGAGGAGAGACGCGCGCACCTCGCCCCGG 840  
Db 992 CGCTGAAAGGCGCGCGCTGCTCGGCGCGAGCGAGGAGCGCGCACCTCGCCCCGG 1051  
QY 841 AGCGTGCAGAGAGACGCGGGCGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCT 900  
Db 1052 AGCGTGCAGAGAGACGCGGGCGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCT 1111  
QY 901 CGCGCTCTCTCGCTCTCAGGGGCGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCT 960  
Db 1112 CGCGCTCTCTCGCTCTCAGGGGCGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCT 1171  
QY 961 GCGCGGAGCGCTCTCGGAGAGAGCA 986  
Db 1172 GCGCGGAGCGCTCTCGGAGAGAGCA 1197

RESULT 2  
AC009108/5  
LOCUS AC009108 168656 bp DNA linear PRI 29-OCT-2002  
DEFINITION Homo sapiens chromosome 16 clone RP11-46309, complete sequence.  
ACCESSION AC009108  
VERSION AC009108.10 GI:24418066  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 168656)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 168656)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 168656)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Submitted (29-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Oct 29, 2002 this sequence version replaced gi:13786306.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos  
National Laboratory  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.2.  
FEATURES  
source  
1..168656  
/organism="Homo sapiens"

/mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="16" /clone="RP11-46309"									
ORIGIN									
Query Match 96.8%; Score 954; DB 9; Length 168656; Best Local Similarity 99.3%; Pred. No. 4, 7e-163; Matches 979; Conservative 0; Mismatches 5; Indels 2; Gaps 2;									
Qy	1	CTGCCAATCCAAATCCAGCGGTTGGCTTTGAATCCATTACACTGGGCCCCCATAATTA	60						
Db	77483	CTGCCAATCCAAATCCAGCGGTTGGC-TTTGAATCCATTACACTGGGCCCCCATAATTA	77425						
Qy	61	GGAAATCTAATATTTCCTTCATCACTCAATTAAGAAAATGCCAGGATCATTTGT	120						
Db	77424	GGAAATCTAATATTTCCTTCATCACTCAATTAAGAAAATGCCAGGATCATTTGT	77365						
Qy	121	ACTTACAAGTCTTTGGGAGAGATATTTTACTCTATTAAATCCATTCTATTTATATTTCA	180						
Db	77364	ACTTACAAGTCTTTGGGAGAGATATTTTACTCTATTAAATCCATTCTATTTATATTTCA	77305						
Qy	181	AATGATTTTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGGCATGTGGGCCAT	240						
Db	77304	AATGATTTTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGGCATGTGGGCCAT	77245						
Qy	241	TCACAAAATGTGATCATATAAATAAATTTTAAAGATATAACTTTTAAAGATTTTCA	300						
Db	77244	TCACAAAATGTGATCATATAAATAAATTTTAAAGATATAACTTTTAAAGATTTTCA	77185						
Qy	301	AGTGAACAGAGATCGCCGCGAGGCGCGGGCGGGGCTTTAGAGCGACGAGATCTCT	360						
Db	77184	AGTGAACAGAGATCGCCGCGAGGCGCGGGGCTTTAGAGCGACGAGATCTCT	77125						
Qy	361	GGCTCTCTCCCGCGATTTGGCGCGGACTCTCTCAGCTCGCGGGTATTTGGCTCAAGT	420						
Db	77124	GGCTCTCTCCCGCGATTTGGCGCGGACTCTCTCAGCTCGCGGGTATTTGGCTCAAGT	77065						
Qy	421	TCGGGAGGGGGCGTGCCCGAGAAAGTAAATCTGCTTTAGCAAGAACTTTTCA	480						
Db	77065	TCGGGAGGGGGCGTGCCCGAGAAAGTAAATCTGCTTTAGCAAGAACTTTTCA	77006						
Qy	481	AATTTTCCCAATCCCTAAAGAGACTTGGCTCTTTTCTGGCTCAGCGGCGAGCG	540						
Db	77005	AATTTTCCCAATCCCTAAAGAGACTTGGCTCTTTTCTGGCTCAGCGGCGAGCG	76946						
Qy	541	CTCGGACCCCGCGCGCTGACCTCGGGGCTGCGGATTCGCTGGGGGCTTGAGAGCCCTC	600						
Db	76945	CTCGGACCCCGCGCGCTGACCTCGGGGCTGCGGATTCGCTGGGGGCTTGAGAGCCCTC	76886						
Qy	601	CTCGGCGCTCTCTCGCGCGGCGAGGTCACCTTGTGTCACGCGCGCGGCTCTCG	660						
Db	76885	CTCGGCGCTCTCTCTCGCGCGGCGAGGTCACCTGCTGTCACGCGCGGCTCTCG	76826						
Qy	661	CTGGGTCGCGCGCGCGCTGCGCGCTGCGCGCGGCTCTGAGCCAGCGAGG	720						
Db	76825	CTGGGTCGCGCGCGCGCTGCGCGCTGCGCGCGGCTCTGAGCCAGCGAGG	76765						
Qy	721	AGCGGGCGCGCGCTGCGCTTGGCGCGCGCGCTTCCAGGATCCGATTCGCGCGGTC	780						
Db	76765	AGCGGGCGCGCGCTGCGCTTGGCGCGCGCGCTTCCAGGATCCGATTCGCGCGGTC	76706						
Qy	781	CGCTGAAAGCGCGCGCTGCTCGGCGCGAGGAGACCGCGCACCTCGCCCGCG	840						
Db	76705	CGCTGAAAGCGCGCGCTGCTCGGCGCGAGGAGACCGCGCACCTCGCCCGCG	76646						
Qy	841	AGGCTGTCAGGAGACCGGGGCGCGCTTCCCGCTTCCCTCTCTCCCTCTCTCTCT	900						
Db	76645	AGGCTGTCAGGAGACCGGGGCGCGCTTCCCGCTTCCCTCTCTCTCTCTCTCTCT	76586						
Qy	901	CGGCT	960						
Db	76585	CGGCT	76526						

Qy 961 GCCGCGAGCGCTCTCGGAGCAGCA 986  
Db 76525 GCCGCGAGCGCTCTCGGAGCAGCA 76500

RESULT 3  
LOCUS AC127554 178416 bp DNA linear ROD 27-NOV-2003  
DEFINITION Mus musculus BAC clone RP24-323K23 from chromosome 8, complete sequence.  
ACCESSION AC127554  
VERSION AC127554.4 GI:33457241  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 178416)  
Cordes, M. and Haglund, K.  
The sequence of Mus musculus BAC clone RP24-323K23  
Unpublished (2001)  
2 (bases 1 to 178416)  
Wilson, R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 178416)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (17-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 178416)  
Wilson, R.K.  
Direct Submission  
Submitted (30-JUL-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 178416)  
Wilson, R.K.  
Direct Submission  
Submitted (06-AUG-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 178416)  
Wilson, R.  
Direct Submission  
Submitted (27-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Aug 6, 2003 this sequence version replaced gi:33342444.  
-----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@wustl.edu  
-----  
Summary Statistics  
Center project name: M\_B80323K23  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see







LOCUS PM212B 869 bp DNA linear STS 29-MAY-2003  
DEFINITION Penicillium marneffei STS, clone pm2a12.b, sequence tagged site.  
ACCESSION AL684671  
VERSION AL684671.1 GI:19337593  
KEYWORDS STS.  
SOURCE Penicillium marneffei  
ORGANISM Penicillium marneffei  
REFERENCE 1  
AUTHORS Yuen,K.Y., Pascal,G., Wong,S.S., Glaser,P., Woo,P.C., Kunst,F.,  
Cai,J.J., Cheung,E.Y., Medigue,C. and Danchin,A.  
TITLE Exploring the penicillium marneffei genome  
JOURNAL Arch. Microbiol. 179 (5), 339-353 (2003)  
MEDLINE 22555073  
REFERENCE 2 (bases 1 to 869)  
AUTHORS Danchin,A. and Pascal,G.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre,  
Dexter HC Man Building 8, Sassoon Road, Pokfulam, Hong Kong  
FEATURES  
source  
1..869  
/organism="Penicillium marneffei"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:37727"  
/clone="pm2a12.b"  
ORIGIN  
Query Match 8.4%; Score 83; DB 11; Length 869;  
Best Local Similarity 48.9%; Pred. No. 3.7e-05;  
Matches 218; Conservative 0; Mismatches 226; Indels 2; Gaps 1;  
QY 520 CTGGGCTCAGCGGGCAGCGCTCGAGACCCCGCGCGTGCACCTCGGGGTGCGCATC 579  
Db 590 CGGCCCCCGGCGGGCCCCCGGGGGCGGCTTCGGCCCGCGCGCCCGCCCGCGNC 531  
QY 580 GTTGGGGGTTGAGAGCTCTCTGCGCCCTCTCTGCGCGGCGCGAGGTCCACCTTGT 639  
Db 530 CCCCCCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 471  
QY 640 CCGAGGCGCGCGCTCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 699  
Db 470 CGGCGCGCGCGCGCTCTCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 411  
QY 700 CGGGTCTCTGAGCAGCAGAGAGCGGGCGCGCGCTG--CGCTTGCGCGCGCGCGCT 757  
Db 410 GGGCGGGCGCGCGCGCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCG 351  
QY 758 CAGGATGCGGATCCGCGCGCGCTCGCTGAAGCGCGCGCGCGCGCGCGCGCGCG 817  
Db 350 CCCCCCG 291  
QY 818 ACAGCGCGCACCTTCGCGCGCGAGAGCTGCAGAGAGCGCGCGCGCGCGCGCG 877  
Db 290 CCG 231  
QY 878 CTCTCTCCCCCTCTGGCTCTCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 937  
Db 230 CG 171  
QY 938 GCGCGCGAGTCCGCGCGCGGGCGCC 963  
Db 170 CG 145  
RESULT 7  
BX664728/c 172457 bp DNA linear HTG 06-NOV-2003  
LOCUS BX664728  
DEFINITION Danio rerio clone CH211-108M10, WORKING DRAFT SEQUENCE, 2 unordered  
pieces.  
ACCESSION BX664728  
VERSION BX664728.4 GI:38201317

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
REFERENCE 1 (bases 1 to 172457)  
AUTHORS McLaren,S.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 7, 2003 this sequence version replaced gi:37805598.  
COMMENT ----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zfsh-help@sanger.ac.uk  
----- Project Information  
Center project name: zc108M10  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 171280 bases at least Q40  
Consensus quality: 171415 bases at least Q30  
Consensus quality: 171619 bases at least Q20  
Insert size: 172357; sum-of-contigs  
Insert size: 177300; 4.4% error; agarose-fp  
Quality coverage: 10.42x in Q20 bases; sum-of-contigs Quality  
coverage: 10.22x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 110971: contig of 110971 bp in length  
\* 110972 111071: gap of 100 bp  
\* 111072 172457: contig of 61366 bp in length.  
FEATURES  
source  
1..172457  
Location/Qualifiers  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-108M10"  
/clone\_lib="CHORI-211"  
misc\_feature  
1..110971  
/note="assembly fragment:01040  
fragment\_chain:1  
clone\_end:T7  
vector\_side:left"  
111072..172457  
/note="assembly fragment:00009  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:right"  
misc\_feature  
111072..172457  
/note="assembly fragment:00009  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:right"  
ORIGIN  
Query Match 8.2%; Score 80.6; DB 2; Length 172457;  
Best Local Similarity 50.8%; Pred. No. 0.00019;  
Matches 217; Conservative 0; Mismatches 209; Indels 1; Gaps 1;  
QY 523 GGCTCAGCGGGCAGCGCTCGGACCCCGCGCGCTGACCTCGGGGCTGCCGATTGCT 582  
Db 110830 GTCTCTGCG 110771  
QY 593 GGGGGCTTGGAGAGCTCTCGCGCCCTCTCTCGCGCGCGCGCGCGCGCGCG 642  
Db 110770 CGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110711





Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 51753)

REFERENCE  
Barron, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL

Direct Submission  
Submitted (17-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 17, 2002 this sequence version replaced gi:21426260.

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information  
Center project name: L27453

Center clone name: 3026\_N\_10

-----

NOTE: This record contains 63 individual

\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1 745: contig of 745 bp in length

\* 746 845: gap of 100 bp

\* 846 1566: contig of 721 bp in length

\* 1567 1666: gap of 100 bp

\* 1667 2371: contig of 705 bp in length

\* 2372 2471: gap of 100 bp

\* 2472 3202: contig of 731 bp in length

\* 3203 3302: gap of 100 bp

\* 3303 4029: contig of 727 bp in length

\* 4030 4129: gap of 100 bp

\* 4130 4826: contig of 697 bp in length

\* 4827 4926: gap of 100 bp

\* 4927 5624: contig of 698 bp in length

\* 5625 5724: gap of 100 bp

\* 5725 6439: contig of 715 bp in length

\* 6440 6539: gap of 100 bp

\* 6540 7262: contig of 723 bp in length

\* 7263 8100: contig of 738 bp in length

\* 8101 8200: gap of 100 bp

\* 8201 8923: contig of 723 bp in length

\* 8924 9023: gap of 100 bp

\* 9024 9747 9846: contig of 723 bp in length

\* 9747 10568: contig of 722 bp in length

\* 10569 10668: gap of 100 bp

\* 10669 11371: contig of 703 bp in length

\* 11372 11471: gap of 100 bp

\* 11472 12184: contig of 713 bp in length

\* 12185 12884: gap of 100 bp

\* 12885 12998 12997: contig of 713 bp in length

\* 12998 13844: contig of 747 bp in length

\* 13845 13944: gap of 100 bp

\* 13945 14668: contig of 724 bp in length

\* 14669 14768: gap of 100 bp

\* 14769 15498: contig of 730 bp in length

\* 15499 15598: gap of 100 bp

\* 15599 16338: contig of 740 bp in length

\* 16339 16438: gap of 100 bp

\* 16439 17194: contig of 756 bp in length

\* 17195 17294: gap of 100 bp

\* 17295 18037: contig of 743 bp in length

\* 18038 18138: gap of 100 bp

\* 18139 18863: contig of 725 bp in length

\* 18863 18962: gap of 100 bp

\* 18963 19640 19739: gap of 100 bp

\* 19640 20459: contig of 720 bp in length

\* 20460 20559: gap of 100 bp

\* 20560 21280: contig of 721 bp in length

\* 21281 21380: gap of 100 bp

\* 21381 22123: contig of 743 bp in length

\* 22124 22223: gap of 100 bp

\* 22224 22957: contig of 734 bp in length

\* 22958 23057: gap of 100 bp

\* 23058 23786: contig of 729 bp in length

\* 23787 23886: gap of 100 bp

\* 23887 24627: contig of 741 bp in length

\* 24628 24727: gap of 100 bp

\* 24728 25456: contig of 729 bp in length

\* 25457 25556: gap of 100 bp

\* 25557 26269: contig of 713 bp in length

\* 26270 26369: gap of 100 bp

\* 26370 27058: contig of 685 bp in length

\* 27059 27158: gap of 100 bp

\* 27159 27881: contig of 723 bp in length

\* 27882 27981: gap of 100 bp

\* 27982 28709: contig of 728 bp in length

\* 28710 28809: gap of 100 bp

\* 28810 29555: contig of 746 bp in length

\* 29556 30384: contig of 725 bp in length

\* 30385 30484: gap of 100 bp

\* 30485 31175: contig of 691 bp in length

\* 31176 31275: gap of 100 bp

\* 31276 31973: contig of 698 bp in length

\* 31974 32073: gap of 100 bp

\* 32074 32790: contig of 717 bp in length

\* 32791 32890: gap of 100 bp

\* 32891 33629: contig of 735 bp in length

```
* 33630 33729: gap of 100 bp
* 33730 34451: contig of 722 bp in length
* 34451 34551: gap of 100 bp
* 34551 35289: contig of 738 bp in length
* 35289 35390: gap of 100 bp
* 35390 36103: contig of 714 bp in length
* 36103 36203: gap of 100 bp
* 36203 36890: contig of 687 bp in length
* 36890 36990: gap of 100 bp
* 36990 37719: contig of 729 bp in length
* 37719 37819: gap of 100 bp
* 37819 38536: contig of 717 bp in length
* 38536 38637: gap of 100 bp
* 38637 39373: contig of 737 bp in length
* 39373 39474: gap of 100 bp
* 39474 40222: contig of 749 bp in length
* 40222 40323: gap of 100 bp
* 40323 41051: contig of 729 bp in length
* 41051 41151: gap of 100 bp
* 41151 41897: contig of 746 bp in length
* 41897 41997: gap of 100 bp
* 41997 42748: contig of 751 bp in length
* 42748 42848: gap of 100 bp
* 42848 43582: contig of 734 bp in length
* 43582 43682: gap of 100 bp
* 43682 44381: contig of 699 bp in length
* 44381 44481: gap of 100 bp
* 44481 45210: contig of 729 bp in length
* 45210 45310: gap of 100 bp
* 45310 46022: contig of 712 bp in length
* 46022 46122: gap of 100 bp
* 46122 46840: contig of 718 bp in length
* 46840 46940: gap of 100 bp

Query Match      8.1%  Score 79.8;  DB 2;  Length 51753;
Best Local Similarity 38.4%  Pred No. 0.00023;
Matches 195;  Conservative 0;  Mismatches 310;  Indels 0;  Gaps 0;

QY 439 CCGAGGAAAGTAAACCTGCTTTCAGCAGAGACTTTTGAAACTTTTCCCAATCCCTA 498
DB 849 CCGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 790
QY 499 AAGGAGACTTGGCTCTTTTCTGGGCTCAGCGGGCAGCGCTCGAGCCCGGCGCT 558
DB 789 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 730
QY 559 GACCTCTGGGGCTCGCGATTCTGCTGGGGGCTTGGAGAGCTCTCTGCGCCCTCTCGCG 618
DB 729 GCGCCCGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
QY 619 GGGCCAGAGGTCCACTTGTGTCCTCCAGCGCGGCGCTCTCGCTGGGTGCGGCGCGCG 678
DB 669 CCGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 610
QY 679 GCCTGCGCGCGCTGCGCGCGCGCGGCTCTGGAGCAGCGAGGAGCGGCGCGCGCTGCG 738
DB 609 GCGCGCGCGCGCGCTGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 550
QY 739 CTTGCGCGGGGCGCGCTTCCAGATGCGGATCGCGCGGTCGCTGGAAGCGCGCGCGCG 798
DB 549 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 490
QY 799 CTGCTCGCGCGGAGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 858
DB 489 NCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
QY 859 GCGCGCGCGCTCGCGCTCGCGCTCTCTCTCGCGCGCTCTCTCGCGCTCTCTCTCTCA 918
DB 429 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 370
QY 919 GGGCGCGCGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 943
DB 369 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345
```

## RESULT 10

AC027483/c

LOCUS

DEFINITION

AC027483

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC027483 Homo sapiens chromosome 3 clone RP11-336E13 map 3, LOW-PASS  
SEQUENCE SAMPLING.  
75144 bp DNA linear HTG 30-MAR-2000

AC027483.1 GI:7342227  
HTG; HTGS\_PHASE0.  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (Bases 1 to 75144)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 3, clone RP11-336E13

2 (Bases 1 to 75144)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,  
Boguslavsky, L., Bouckgeater, B., Brown, A., Burkett, G.,  
Campopiano, A., Castie, A., Choepel, Y., Colangelo, W., Collins, S.,  
Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehotzky, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testa, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L8413  
Center clone name: 336\_E\_13

-----  
\* NOTE: This record contains 88 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 742: contig of 742 bp in length  
743 842: gap of 100 bp  
843 1604: contig of 762 bp in length  
1605 1704: gap of 100 bp  
1705 2466: contig of 762 bp in length  
2467 2566: gap of 100 bp  
2567 3325: contig of 759 bp in length

```
* 3326 3425: gap of 100 bp
* 3426 4178: contig of 753 bp in length
* 4179 4278: gap of 100 bp
* 4279 5034: contig of 756 bp in length
* 5035 5134: gap of 100 bp
* 5135 5884: contig of 750 bp in length
* 5885 5984: gap of 100 bp
* 5985 6758: contig of 774 bp in length
* 6759 6858: gap of 100 bp
* 6859 7587: contig of 723 bp in length
* 7588 8425: contig of 738 bp in length
* 8426 8525: gap of 100 bp
* 8526 9282: contig of 757 bp in length
* 9283 10126: gap of 100 bp
* 10127 10226: contig of 744 bp in length
* 10227 10998: contig of 772 bp in length
* 10999 11098: gap of 100 bp
* 11099 11838: contig of 740 bp in length
* 11839 11938: gap of 100 bp
* 11939 12704: contig of 766 bp in length
* 12705 12804: gap of 100 bp
* 12805 13539: contig of 735 bp in length
* 13540 14389: contig of 750 bp in length
* 14390 14489: gap of 100 bp
* 14490 15233: contig of 744 bp in length
* 15234 15333: gap of 100 bp
* 15334 16094: contig of 761 bp in length
* 16095 16194: gap of 100 bp
* 16195 16894: contig of 700 bp in length
* 16895 16994: gap of 100 bp
* 16995 17751: contig of 757 bp in length
* 17752 17851: gap of 100 bp
* 17852 18629: contig of 778 bp in length
* 18630 18729: gap of 100 bp
* 18730 19480: contig of 751 bp in length
* 19481 19580: gap of 100 bp
* 19581 20360: contig of 780 bp in length
* 20361 20460: gap of 100 bp
* 20461 21218: contig of 758 bp in length
* 21219 21318: gap of 100 bp
* 21319 22100: contig of 782 bp in length
* 22101 22200: gap of 100 bp
* 22201 22945: contig of 745 bp in length
* 22946 23045: gap of 100 bp
* 23046 23808: contig of 763 bp in length
* 23809 23908: gap of 100 bp
* 23909 24551: contig of 743 bp in length
* 24552 24751: gap of 100 bp
* 24752 25005: contig of 754 bp in length
* 25006 25605: gap of 100 bp
* 25606 26366: contig of 761 bp in length
* 26367 26466: gap of 100 bp
* 26467 27240: contig of 774 bp in length
* 27241 27340: gap of 100 bp
* 27341 28101: contig of 761 bp in length
* 28102 28201: gap of 100 bp
* 28202 28983: contig of 782 bp in length
* 28984 29084: gap of 100 bp
* 29084 29789: contig of 706 bp in length
* 29790 29889: gap of 100 bp
* 29890 30647: contig of 758 bp in length
* 30648 30747: gap of 100 bp
* 30748 31499: contig of 752 bp in length
* 31500 31599: gap of 100 bp
* 31600 32356: contig of 757 bp in length
* 32357 32456: gap of 100 bp
* 32457 33205: contig of 749 bp in length
* 33206 33305: gap of 100 bp
* 33306 34052: contig of 747 bp in length
* 34053 34152: gap of 100 bp

* 34153 34912: contig of 760 bp in length
* 34913 35012: gap of 100 bp
* 35013 35766: contig of 754 bp in length
* 35767 35866: gap of 100 bp
* 35867 36522: contig of 786 bp in length
* 36523 36752: gap of 100 bp
* 36753 37528: contig of 776 bp in length
* 37529 37628: gap of 100 bp
* 37629 38395: contig of 767 bp in length
* 38396 38495: gap of 100 bp
* 38496 39267: contig of 772 bp in length
* 39268 39367: gap of 100 bp
* 39368 40118: contig of 751 bp in length
* 40119 40218: gap of 100 bp
* 40219 40969: contig of 751 bp in length
* 40970 41069: gap of 100 bp
* 41070 41794: contig of 725 bp in length
* 41795 41894: gap of 100 bp
* 41895 42620: contig of 726 bp in length
* 42621 42720: gap of 100 bp
* 42721 43486: contig of 766 bp in length
* 43487 43586: gap of 100 bp
* 43587 44314: contig of 728 bp in length
* 44315 44414: gap of 100 bp
* 44415 45155: contig of 741 bp in length
* 45156 45255: gap of 100 bp
* 45256 46007: contig of 752 bp in length
* 46008 46107: gap of 100 bp
* 46108 46846: contig of 739 bp in length
* 46847 46946: gap of 100 bp
* 46947 47707: contig of 761 bp in length
* 47708 47807: gap of 100 bp
* 47808 48564: contig of 757 bp in length
* 48565 48664: gap of 100 bp
* 48665 49424: contig of 760 bp in length
* 49425 49524: gap of 100 bp
* 49525 50247: contig of 723 bp in length
* 50248 50347: gap of 100 bp
* 50348 51106: contig of 759 bp in length
* 51107 51206: gap of 100 bp
* 51207 51956: contig of 750 bp in length
* 51957 52056: gap of 100 bp
* 52057 52803: contig of 747 bp in length
* 52804 52903: gap of 100 bp
* 52904 53669: contig of 766 bp in length
* 53670 53769: gap of 100 bp
* 53770 54469: contig of 700 bp in length
* 54470 54569: gap of 100 bp
* 54570 55301: contig of 732 bp in length
* 55302 55401: gap of 100 bp
* 55402 56192: contig of 791 bp in length
* 56193 56292: gap of 100 bp
* 56293 57046: contig of 754 bp in length
* 57047 57146: gap of 100 bp
* 57147 57923: contig of 777 bp in length
* 57924 58023: gap of 100 bp
* 58024 58803: contig of 780 bp in length
* 58804 58903: gap of 100 bp

Query Match 8.1%; Score 79.8; DB 2; Length 75144;
Best Local Similarity 50.7%; Pred. No. 0.0024;
Matches 192; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

Qy 598 CTCCTGCGGCGCTCTCGCGCGGCGGAGGTCACCTGTGTCCTCCAGCGCGCGGCTCT 657
Db 15866 CCCCCACCCCGAGCGCCCCCCCCCGGAGCGCCCCCCCCCGGCGCGCGCGCTCT 16807
Qy 658 CCGCTGGGTCCGCGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCGCGGCTCTTGAGCGCAGG 717
Db 16806 CCGAGGCCACGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 16747
Qy 718 AGGAGCGGCGCGCGCGCTGCGCGCTTCCCGGCGCGCGCGCTCCAGATCGCGCGCG 777
```





\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1	817: contig of 817 bp in length	30068	30167: gap of unknown length
*	917: gap of unknown length	*	30991: contig of 824 bp in length
*	818	*	31091: gap of unknown length
*	918	*	31900: contig of 809 bp in length
*	1076	*	32000: gap of unknown length
*	1176	*	33585: contig of 1585 bp in length
*	1972	*	33685: gap of unknown length
*	2072	*	34444: contig of 759 bp in length
*	2856	*	34544: gap of unknown length
*	2856: gap of unknown length	*	35355: contig of 811 bp in length
*	3751: contig of 796 bp in length	*	35455: gap of unknown length
*	3851: gap of unknown length	*	36217: contig of 762 bp in length
*	4649: gap of 798 bp in length	*	36317: gap of unknown length
*	4749: gap of unknown length	*	37149: contig of 832 bp in length
*	5543: contig of 794 bp in length	*	37249: gap of unknown length
*	5643: gap of unknown length	*	38010: contig of 761 bp in length
*	5644: contig of 805 bp in length	*	38110: gap of unknown length
*	5648: gap of unknown length	*	38919: contig of 809 bp in length
*	5649: gap of unknown length	*	39019: gap of unknown length
*	7341	*	39795: contig of 776 bp in length
*	7404: gap of unknown length	*	39895: gap of unknown length
*	7605	*	40711: contig of 816 bp in length
*	7704: gap of unknown length	*	40811: gap of unknown length
*	8512: contig of 808 bp in length	*	41600: contig of 789 bp in length
*	8612: gap of unknown length	*	41700: gap of unknown length
*	9415: contig of 803 bp in length	*	42517: contig of 817 bp in length
*	9515: gap of unknown length	*	42617: gap of unknown length
*	10302: contig of 787 bp in length	*	43398: contig of 781 bp in length
*	10303: gap of unknown length	*	43499: gap of unknown length
*	11207: contig of 805 bp in length	*	44309: contig of 811 bp in length
*	11307: gap of unknown length	*	44409: gap of unknown length
*	12105: contig of 798 bp in length	*	45216: contig of 807 bp in length
*	12205: gap of unknown length	*	45316: gap of unknown length
*	13003: contig of 798 bp in length	*	46129: contig of 813 bp in length
*	13103: gap of unknown length	*	46229: gap of unknown length
*	13104	*	47033: contig of 804 bp in length
*	13902: contig of 799 bp in length	*	47133: gap of unknown length
*	14002: gap of unknown length	*	47943: contig of 810 bp in length
*	14805: contig of 803 bp in length	*	48043: gap of unknown length
*	14905: gap of unknown length	*	48544: contig of 811 bp in length
*	15700: contig of 795 bp in length	*	48954: gap of unknown length
*	15701	*	48955: contig of 811 bp in length
*	15800: gap of unknown length	*	49655: gap of unknown length
*	16609: contig of 809 bp in length	*	49865: gap of unknown length
*	16709: gap of unknown length	*	50670: contig of 805 bp in length
*	16710	*	50770: gap of unknown length
*	17506: contig of 797 bp in length	*	51583: contig of 813 bp in length
*	17606: gap of unknown length	*	51683: gap of unknown length
*	18421: contig of 815 bp in length	*	52496: contig of 813 bp in length
*	18521: gap of unknown length	*	52596: gap of unknown length
*	18621: contig of 1341 bp in length	*	53417: contig of 821 bp in length
*	19662: gap of unknown length	*	53517: gap of unknown length
*	19863	*	54323: contig of 806 bp in length
*	20768	*	54423: gap of unknown length
*	20968	*	55237: contig of 814 bp in length
*	21667: gap of unknown length	*	55337: gap of unknown length
*	21767	*	56158: contig of 821 bp in length
*	22572	*	56258: gap of unknown length
*	22671: gap of unknown length	*	57065: contig of 807 bp in length
*	22790: contig of 119 bp in length	*	57165: gap of unknown length
*	22890: gap of unknown length	*	57954: contig of 789 bp in length
*	23702: contig of 812 bp in length	*	58054: gap of unknown length
*	23802: gap of unknown length	*	58973: contig of 919 bp in length
*	24704: contig of 802 bp in length	*	59073: gap of unknown length
*	25512: contig of 808 bp in length	*	59888: contig of 815 bp in length
*	25612: gap of unknown length	*	59988: gap of unknown length
*	26424: contig of 812 bp in length	*	60795: contig of 807 bp in length
*	26524: gap of unknown length	*	60895: gap of unknown length
*	26525	*	61705: contig of 810 bp in length
*	27329: contig of 805 bp in length	*	61805: gap of unknown length
*	27429: gap of unknown length	*	62606: contig of 801 bp in length
*	28237: contig of 808 bp in length	*	62706: gap of unknown length
*	28337: gap of unknown length	*	63517: contig of 811 bp in length
*	29155: contig of 818 bp in length	*	63617: gap of unknown length
*	29255: gap of unknown length	*	
*	30067: contig of 812 bp in length	*	

JOURNAL Submitted (16-JAN-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
COMMENT On Jan 16, 2003 this sequence version replaced gi:28665814.  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
----- Project Information -----  
Center project name: M.B0444115  
----- Summary Statistics -----  
Sequencing vector: M13, 0%  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 221817 bases at least Q40  
Consensus quality: 22474 bases at least Q30  
Consensus quality: 223020 bases at least Q20  
Insert size: 22000; agarose-fp  
Insert size: 225204; sum-of-contigs  
Quality coverage: 11.88 in Q20 bases; agarose-fp  
Quality coverage: 11.64 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

Query Match 8.0%; Score 79.2; DB 2; Length 85434;  
Best Local Similarity 48.7%; Pred. No. 0.00032;  
Matches 219; Conservative 0; Mismatches 229; Indels 2; Gaps 1;

QY 506 CTTGGCTCTTTTCTGGGCTACGCGGCGAGCGCTCGACCCCGCGCTGACCCCTC 565  
DB 70872 CTTNGCCCGCTTTTNGGNGNCCAAACCGCCNCGNGNNGCCNTGCGCGCGCGCGCCGC 70931  
QY 566 GGGGCTGCCGATTCGCTGGGGCTTGGAGAGCTCTCGGCCCTCTCTCGCGCGCGCGA 625  
DB 70932 CGCCCCCGCGCGCGCGCTTTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 70991  
QY 626 GGTCTCACCTTGTCTCCAGCGCGCGCGCTCTCCGCTGGGTCTCGCGCGCGCGCGCTGCC 685  
DB 70992 GGTCTCAATACG 71051  
QY 686 CGCGCTCGCGCGCGCGCTCTGGAGCAGCAGGAGCGCGCGCGCGCGCGCTGCGCTGCC 745  
DB 71052 CG 71111  
QY 746 GGGCGCGCGCGCTCAGATCCGATCCGCGCGCGCGCTCGCTGAAAGCGCGCGCGCGCTCG 805  
DB 71112 CG 71169  
QY 806 GCGCGAGCAGCAGCAGCG 865  
DB 71170 GCG 71229  
QY 866 CTTCCCGCTCCCT 925  
DB 71230 CG 71289  
QY 926 CTTCTGCTCCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 955  
DB 71290 CCG 71319

## RESULT 14

AC138109 224777 bp DNA linear HTG 16-JAN-2003  
LOCUS Mus musculus chromosome UNK clone RP24-444115, WORKING DRAFT  
DEFINITION SEQUENCE, 8 unordered pieces.  
ACCESSION AC138109  
VERSION AC138109.2 GI:27764789  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 224777)  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE The sequenced of Mus musculus clone  
JOURNAL Unpublished  
2 (bases 1 to 224777)  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (13-DEC-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 224777)  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission

## FEATURES

source  
1. 224777  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP24-444115"  
misc\_feature  
1. 2636  
/note="assembly\_name:Contig16"  
misc\_feature  
2737..7027  
/note="assembly\_name:Contig17"  
misc\_feature  
7128..13377  
/note="assembly\_name:Contig18"  
misc\_feature  
13478..22867  
/note="assembly\_name:Contig19"  
misc\_feature  
22968..33389  
/note="assembly\_name:Contig20"  
misc\_feature  
33490..47420  
/note="assembly\_name:Contig21"  
misc\_feature  
47521..107676  
/note="assembly\_name:Contig22"  
misc\_feature  
107777..224777  
/note="assembly\_name:Contig23"

## ORIGIN







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:55:14 ; Search time 1.77026 Seconds  
(without alignments)  
2821.370 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_223\_231

Perfect score: 9

Sequence: 1 acaatgtt 9

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	100.0	17	1	US-08-758-306-137
C 2	9	100.0	17	1	US-08-758-306-139
C 3	9	100.0	17	1	US-08-758-306-1109
C 4	9	100.0	17	1	US-08-758-306-1111
C 5	9	100.0	17	4	US-08-584-040-5500
C 6	9	100.0	17	4	US-08-584-040-5501
C 7	9	100.0	17	4	US-09-371-772B-2391
C 8	9	100.0	17	4	US-09-371-772B-2392
C 9	9	100.0	20	4	US-09-780-175-137
C 10	9	100.0	20	4	US-09-780-175-138
C 11	9	100.0	22	1	US-08-512-681-19
C 12	9	100.0	22	2	US-08-822-028-35
C 13	9	100.0	22	2	US-08-743-637B-240
C 14	9	100.0	22	3	US-08-479-285-35
C 15	9	100.0	22	3	US-09-503-653A-35
C 16	9	100.0	22	3	US-09-084-220-10
C 17	9	100.0	26	4	US-09-199-542B-85
C 18	9	100.0	29	3	US-09-400-046-5
C 19	9	100.0	30	4	US-09-438-268-34
C 20	9	100.0	30	4	US-09-438-268-35
C 21	9	100.0	35	3	US-09-232-477-14
C 22	9	100.0	35	4	US-09-784-982-14
C 23	9	100.0	40	2	US-08-882-083-4
C 24	9	100.0	40	2	US-08-558-107-4
C 25	9	100.0	40	3	US-09-243-539-4
C 26	9	100.0	41	3	US-08-448-194-54
C 27	9	100.0	41	3	US-08-441-935-9

Sequence 9, Appli  
Sequence 54, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 635, App  
Sequence 1408, Ap  
Sequence 2684, Ap  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 30, Appli  
Sequence 2446, Ap  
Sequence 1720, Ap  
Sequence 2624, Ap  
Sequence 130, App  
Sequence 122, App

28 9 100.0 41 3 US-08-441-943-9  
29 9 100.0 41 4 US-08-867-921-54  
30 9 100.0 46 3 US-08-258-2807B-17  
31 9 100.0 46 3 US-08-368-704C-17  
32 9 100.0 47 4 US-09-671-317-635  
33 9 100.0 47 4 US-09-422-978-1408  
34 9 100.0 47 4 US-08-422-978-2684  
35 9 100.0 48 2 US-08-882-083-3  
36 9 100.0 48 2 US-08-558-107-3  
37 9 100.0 48 3 US-09-243-539-3  
38 9 100.0 49 3 US-08-441-935-7  
39 9 100.0 49 3 US-08-441-943-7  
40 9 100.0 50 4 US-09-849-069-30  
41 9 100.0 53 4 US-08-956-171E-2446  
42 9 100.0 55 4 US-08-956-171E-1720  
43 9 100.0 55 4 US-08-956-171E-2624  
44 9 100.0 71 1 US-08-257-073-130  
45 9 100.0 72 1 US-08-413-118-122

#### ALIGNMENTS

#### RESULT 1

US-08-758-306-137/c  
; Sequence 137, Application US/08758306  
; Patent No. 5807743  
; GENERAL INFORMATION:  
; APPLICANT: McSwiggan, Dan T.  
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
; TITLE OF INVENTION: TREATMENT OF DISEASES  
; TITLE OF INVENTION: ASSOCIATED WITH  
; TITLE OF INVENTION: INTERLEUKIN-2 RECEPTOR  
; TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION  
; NUMBER OF SEQUENCES: 1379  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Fastseq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/758,306  
; FILING DATE: December 3, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 212/132  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 137:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-758-306-137

Query Match 100.0%; Score 9; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
Db 12 ACAATGTT 4

## RESULT 2

US-08-758-306-139/c  
; Sequence 139, Application US/08758306  
; Patent No. 5807743  
; GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: McSwiggen, James A.  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TREATMENT OF DISEASES  
TITLE OF INVENTION: ASSOCIATED WITH  
INTERLEUKIN-2 RECEPTOR  
TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION  
NUMBER OF SEQUENCES: 1379  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/758,306  
FILING DATE: December 3, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 212/132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 139:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-758-306-139

Query Match 100.0%; Score 9; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
Db 11 ACAATGTT 3

## RESULT 3

US-08-758-306-1109/c  
; Sequence 1109, Application US/08758306  
; Patent No. 5807743  
; GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: McSwiggen, James A.  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TREATMENT OF DISEASES  
TITLE OF INVENTION: ASSOCIATED WITH  
INTERLEUKIN-2 RECEPTOR  
TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION  
NUMBER OF SEQUENCES: 1379  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/758,306  
FILING DATE: December 3, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 212/132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1109:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-758-306-1109

Query Match 100.0%; Score 9; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
Db 12 ACAATGTT 4

## RESULT 4

US-08-758-306-1111/c  
; Sequence 1111, Application US/08758306  
; Patent No. 5807743  
; GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: McSwiggen, James A.  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TREATMENT OF DISEASES  
TITLE OF INVENTION: ASSOCIATED WITH  
INTERLEUKIN-2 RECEPTOR  
TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION  
NUMBER OF SEQUENCES: 1379  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California

```

; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,306
; FILING DATE: December 3, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-758-306-1111

```

```

Query Match 100.0%; Score 9; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ACAAATGTT 9
   |||||
DB 11 ACAAATGTT 3

```

```

RESULT 5
US-08-584-040-5500/c
; Sequence 5500, Application US/08584040
; Patent No. 6346398
; GENERAL INFORMATION:
; APPLICANT: Pavco, Pamela
; APPLICANT: McSwiggen, James
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: TREATMENT OF DISEASES OR
; TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
; TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 8502
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/584,040
; FILING DATE: January 11, 1996
; CLASSIFICATION: 514

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,974
; FILING DATE: October 26, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 5500:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-584-040-5500

```

```

Query Match 100.0%; Score 9; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ACAAATGTT 9
   |||||
DB 12 ACAAATGTT 4

```

```

RESULT 6
US-08-584-040-5501/c
; Sequence 5501, Application US/08584040
; Patent No. 6346398
; GENERAL INFORMATION:
; APPLICANT: Pavco, Pamela
; APPLICANT: McSwiggen, James
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: TREATMENT OF DISEASES OR
; TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
; TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 8502
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/584,040
; FILING DATE: January 11, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,974
; FILING DATE: October 26, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 5501:

```

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 17 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
US-08-584-040-5501

Query Match      100.0%; Score 9; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAATGTT 9
    |||||
Db 11 ACAAAATGTT 3

RESULT 7
US-09-371-772B-2391/c
; Sequence 2391, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MEH800.876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2391
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Mus sp.
US-09-371-772B-2391

Query Match      100.0%; Score 9; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAATGTT 9
    |||||
Db 12 ACAAAATGTT 4

RESULT 8
US-09-371-772B-2392/c
; Sequence 2392, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MEH800.876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
```

```
;
; SEQ ID NO 2392
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Mus sp.
US-09-371-772B-2392

Query Match      100.0%; Score 9; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAATGTT 9
    |||||
Db 11 ACAAAATGTT 3

RESULT 9
US-09-780-175-137/c
; Sequence 137, Application US/09780175
; Patent No. 6440738
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-BETA EXPRESSION
; FILE REFERENCE: RTS-0164
; CURRENT APPLICATION NUMBER: US/09/780,175
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 137
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-780-175-137

Query Match      100.0%; Score 9; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAATGTT 9
    |||||
Db 9 ACAAAATGTT 1

RESULT 10
US-09-780-175-138/c
; Sequence 138, Application US/09780175
; Patent No. 6440738
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-BETA EXPRESSION
; FILE REFERENCE: RTS-0164
; CURRENT APPLICATION NUMBER: US/09/780,175
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 138
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-780-175-138

Query Match      100.0%; Score 9; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAATGTT 9
    |||||
Db 20 ACAAAATGTT 12
```

RESULT 11  
US-08-512-681-19/c  
Sequence 19, Application US/08512681  
Patent No. 5795976  
GENERAL INFORMATION:  
APPLICANT: Oefner, Peter J.  
APPLICANT: Underhill, Peter A.  
TITLE OF INVENTION: Detection of DNA Heteroduplex Molecules  
TITLE OF INVENTION: By Denaturing High Performance Liquid Chromatography and  
TITLE OF INVENTION: Methods for Comparative Sequencing  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/512,681  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, Susan T.  
REGISTRATION NUMBER: 38,443  
REFERENCE/DOCKET NUMBER: 8600-0155  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: DYS234 REVERSE PRIMER  
US-08-512-681-19

Query Match 100.0%; Score 9; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9  
|||||  
Db 14 ACAAATGTT 6

RESULT 12  
US-08-822-028-35  
Sequence 35, Application US/08822028  
Patent No. 5993813  
GENERAL INFORMATION:  
APPLICANT: MEZES, PETER S  
APPLICANT: GOURLIE, BRIAN B  
APPLICANT: RIXON, MARK W  
APPLICANT: ANDERSON, WH KERR  
APPLICANT: KAPLAN, DONALD A  
APPLICANT: SCHOLOW, JEFFREY  
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,  
TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DUANE C ULMER  
STREET: P.O. BOX 1967  
CITY: MIDLAND  
STATE: MICHIGAN  
COUNTRY: USA

ZIP: 48641-1967  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/040,687  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: ULMER, DUANE C  
REGISTRATION NUMBER: 34,941  
REFERENCE/DOCKET NUMBER: C-37,075C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 636-8104  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-822-028-35

Query Match 100.0%; Score 9; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9  
|||||  
Db 12 ACAAATGTT 20

RESULT 13  
US-08-743-637B-240/c  
Sequence 240, Application US/08743637B  
Patent No. 5994066  
GENERAL INFORMATION:  
APPLICANT: BERGERON, Michel G.  
APPLICANT: PICARD, Francois J.  
APPLICANT: OUELLETTE, Marc  
APPLICANT: ROY, Paul H.  
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA  
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND  
TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED  
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS  
NUMBER OF SEQUENCES: 273  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: QUARLES & BRADY  
STREET: 411 EAST WISCONSIN AVENUE  
CITY: MILWAUKEE  
STATE: WISCONSIN  
COUNTRY: USA  
ZIP: 53202-4497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,637B  
FILING DATE: 04-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/526,840  
FILING DATE: 11-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BAKER, Jean C.  
REGISTRATION NUMBER: 35,433  
REFERENCE/DOCKET NUMBER: 850586.90012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (414) 277-5000  
TELEFAX: (414) 277-5591  
INFORMATION FOR SEQ ID NO: 240:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-743-637B-240

Query Match 100.0%; Score 9; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
Db 13 ACAAATGTT 5

## RESULT 14

US-08-479-285-35  
Sequence 35, Application US/08479285  
Patent No. 6207815  
GENERAL INFORMATION:  
APPLICANT: MEZES, PETER S  
APPLICANT: GOURLIE, BRIAN B  
APPLICANT: RIXON, MARK W  
APPLICANT: ANDERSON, WH KERR  
APPLICANT: KAPLAN, DONALD A  
APPLICANT: SCHOLOM, JEFFREY  
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,  
MODIFIED ANTIBODIES FOR CANCER TREATMENT  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DUANE C ULMER  
STREET: P.O. BOX 1967  
CITY: MIDLAND  
STATE: MICHIGAN  
COUNTRY: USA  
ZIP: 48641-1967  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,285  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/040687  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: ULMER, DUANE C  
REGISTRATION NUMBER: 34,941  
REFERENCE/DOCKET NUMBER: C-37,075C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 636-8104  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-479-285-35

Query Match 100.0%; Score 9; DB 3; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9

Db 12 ACAAATGTT 20

## RESULT 15

US-09-503-653A-35  
Sequence 35, Application US/09503653A  
Patent No. 6641999  
GENERAL INFORMATION:  
APPLICANT: MEZES, PETER S  
APPLICANT: GOURLIE, BRIAN B  
APPLICANT: RIXON, MARK W  
APPLICANT: ANDERSON, WH KERR  
APPLICANT: KAPLAN, DONALD A  
APPLICANT: SCHLOM, JEFFREY  
TITLE OF INVENTION: Probing Method for Identifying Antibodies  
FILE REFERENCE: 37075H-CIP1  
CURRENT APPLICATION NUMBER: US/09/503.653A  
CURRENT FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: US 08/040,687  
PRIOR FILING DATE: 1993-03-31  
PRIOR APPLICATION NUMBER: US 07/424,362  
PRIOR FILING DATE: 1989-10-19  
PRIOR APPLICATION NUMBER: US 07/261,942  
PRIOR FILING DATE: 1988-10-24  
PRIOR APPLICATION NUMBER: US 07/259,943  
PRIOR FILING DATE: 1988-10-19  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: MICROSOFT Word 97 SR-2  
SEQ ID NO 35  
LENGTH: 22  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..22  
OTHER INFORMATION: Oligo CC83 L intron(-)  
US-09-503-653A-35

Query Match 100.0%; Score 9; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
Db 12 ACAAATGTT 20

Search completed: March 25, 2004, 15:34:20  
Job time : 2.77026 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 14.7262 Seconds  
(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375

Perfect score: 17

Sequence: 1 tctgaaggaataaata 17

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2459946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications NA:\*
- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
  - 10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
  - 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
  - 13: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
  - 14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
  - 15: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
  - 16: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
  - 17: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
  - 18: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	100.0	962	15	US-10-027-632-9718
2	17	100.0	658	9	US-09-963-285-1
3	16	94.1	493	12	US-10-027-632-9718
4	16	94.1	2523	12	US-10-027-632-9718
5	16	94.1	21619	10	US-09-764-891-10003
6	16	94.1	21619	14	US-10-205-428-977
7	15.4	90.6	50	15	US-10-131-827-1138
8	15.4	90.6	359	12	US-10-027-632-9718
9	15.4	90.6	386	10	US-10-027-632-9718
10	15.4	90.6	493	12	US-10-027-632-9718
11	15.4	90.6	589	15	US-10-027-632-9718
12	15.4	90.6	589	15	US-10-027-632-9718
13	15.4	90.6	606	12	US-10-027-632-9718
14	15.4	90.6	616	12	US-10-027-632-9718
15	15.4	90.6	676	15	US-10-027-632-9718

c	16	15.4	90.6	820	12	US-10-027-632-9718	Sequence 37752, A
c	17	15.4	90.6	831	12	US-10-027-632-9718	Sequence 43641, A
	18	15.4	90.6	1228	15	US-10-027-632-9718	Sequence 256262, A
	19	15.4	90.6	1228	15	US-10-027-632-9718	Sequence 256263, A
	20	15.4	90.6	1228	15	US-10-027-632-9718	Sequence 256264, A
	21	15.4	90.6	1555	14	US-10-066-998-95	Sequence 95, Appl
	22	15.4	90.6	1780	15	US-10-087-080-24	Sequence 24, Appl
	23	15.4	90.6	2336	12	US-10-229-345-12	Sequence 12, Appl
	24	15.4	90.6	2336	12	US-10-374-177-12	Sequence 17, Appl
	25	15.4	90.6	2336	15	US-10-341-434-17	Sequence 12, Appl
	26	15.4	90.6	2336	15	US-10-087-080-27	Sequence 27, Appl
	27	15.4	90.6	5219	15	US-10-062-674-2121	Sequence 2121, Ap
	28	15.4	90.6	6021	9	US-09-963-285-5	Sequence 5, Appl
	29	15.4	90.6	6052	12	US-10-221-613-351	Sequence 351, App
	30	15.4	90.6	7133	14	US-10-198-846-13766	Sequence 13766, A
	31	15.4	90.6	8197	14	US-10-240-485-68	Sequence 68, Appl
	32	15.4	90.6	16281	9	US-09-764-847-1367	Sequence 1367, Ap
	33	15.4	90.6	16281	14	US-10-092-154-1367	Sequence 1368, Ap
	34	15.4	90.6	16285	9	US-09-764-847-1368	Sequence 1369, Ap
	35	15.4	90.6	16285	9	US-09-764-847-1369	Sequence 1368, Ap
	36	15.4	90.6	16285	14	US-10-092-154-1368	Sequence 1369, Ap
	37	15.4	90.6	16285	14	US-10-092-154-1369	Sequence 1369, Ap
	38	15.4	90.6	17294	14	US-10-311-455-959	Sequence 959, Appl
	39	15.4	90.6	52216	9	US-09-747-810-1	Sequence 1, Appl
	40	15.4	90.6	180557	13	US-10-003-806-6	Sequence 9, Appl
	41	15.4	90.6	180557	13	US-10-003-806-9	Sequence 1, Appl
	42	15.4	90.6	186510	14	US-10-043-715-1	Sequence 8323, Ap
c	43	15	88.2	276	12	US-10-073-644C-1	Sequence 1, Appl
	44	15	88.2	357	14	US-10-073-644C-1	Sequence 234237,
	45	15	88.2	581	15	US-10-027-632-234237	

ALIGNMENTS

RESULT 1  
US-10-027-632-9718  
Sequence 9718, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027.632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 9718  
LENGTH: 962  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-9718

Query Match 100.0%; Score 17; DB 15; Length 962;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1. TGTGAAGGAATAAATA 17

```
Db      727  TGTGGAAGGAATAATA 743
|||||
|||||

RESULT 2
US-09-963-285-1
; Sequence 1, Application US/09963285
; Patent No. US20020090707A1
; GENERAL INFORMATION:
; APPLICANT: Enerbeck, Sven
; APPLICANT: Krook, Katarina
; APPLICANT: Rondahl, Lena
; APPLICANT: Wasserman, Wytch
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-042001
; CURRENT APPLICATION NUMBER: US/09/963,285
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: SE 0004102-0
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/238,897
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: SE 0003435-5
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2235)...(3737)
US-09-963-285-1

Query Match      100.0%; Score 17; DB 9; Length 6458;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TGTGGAAGGAATAATA 17
|||||
|||||

Db      359  TGTGGAAGGAATAATA 375
|||||
|||||

RESULT 3
US-10-424-599-107702
; Sequence 107702, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 107702
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(493)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68271C.1
US-10-424-599-107702

Query Match      94.1%; Score 16; DB 12; Length 493;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      2  GTGGAAGGAATAATA 17
|||||
|||||

Db      233  GTGGAAGGAATAATA 248
|||||
|||||

RESULT 4
US-10-424-599-82796
; Sequence 82796, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 82796
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45782C.1
US-10-424-599-82796

Query Match      94.1%; Score 16; DB 12; Length 2523;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2  GTGGAAGGAATAATA 17
|||||
|||||

Db      502  GTGGAAGGAATAATA 617
|||||
|||||

RESULT 5
US-09-764-891-10003/C
; Sequence 10003, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10003
; LENGTH: 21619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10003

Query Match      94.1%; Score 16; DB 10; Length 21619;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TGTGGAAGGAATAATA 16
|||||
|||||

Db      5310  TGTGGAAGGAATAATA 5295
|||||
|||||

RESULT 6
US-10-205-428-977/C
; Sequence 977, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```

; FILE REFERENCE: PAL17C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 977
; LENGTH: 21619
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-428-977

Query Match          94.1%; Score 16; DB 14; Length 21619;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGTGGAAGGAATAAAT 16
      |||
DB      5310 TGTGGAAGGAATAAAT 5295

RESULT 7
US-10-131-827-1138
; Sequence 1138, Application US/10/131,827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1138
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-131-827-1138

Query Match          90.6%; Score 15.4; DB 15; Length 50;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TGTGGAAGGAATAAAT 17
      |||
DB      5310 TGTGGAAGGAATAAAT 5295

FILE REFERENCE: PAL17C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 977
; LENGTH: 21619
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-428-977

Query Match          94.1%; Score 16; DB 14; Length 21619;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGTGGAAGGAATAAAT 16
      |||
DB      5310 TGTGGAAGGAATAAAT 5295

RESULT 8
US-10-424-599-64036
; Sequence 64036, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 64036
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT3847_28838C.1
; US-10-424-599-64036

Query Match          90.6%; Score 15.4; DB 12; Length 359;
Best Local Similarity 94.1%; Pred. No. 8.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TGTGGAAGGAATAAATA 17
      |||
DB      59 TGTGGAAGGAATAAATA 75

RESULT 9
US-03-814-353-17187
; Sequence 17187, Application US/09814353
; Publication No. US20030185831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17187
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-03-814-353-17187

Query Match          90.6%; Score 15.4; DB 10; Length 386;
Best Local Similarity 94.1%; Pred. No. 8.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```



US-10-424-599-141232

Query Match 90.6%; Score 15.4; DB 12; Length 606;  
Best Local Similarity 94.1%; Pred. No. 8.5e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGGAAGGATAAATA 17  
|||||  
DB 513 TGTGGAAGGATAAATA 529

## RESULT 14

US-10-424-599-63676/c  
; Sequence 63676, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 63676  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_28513C.1  
US-10-424-599-63676

Query Match 90.6%; Score 15.4; DB 12; Length 616;  
Best Local Similarity 94.1%; Pred. No. 8.5e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGGAAGGATAAATA 17  
|||||  
DB 264 TGTGGAAGGATAAATA 248

## RESULT 15

US-10-027-632-148185/c  
; Sequence 148185, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 148185  
; LENGTH: 676  
; TYPE: DNA

; ORGANISM: Human  
US-10-027-632-148185

Query Match 90.6%; Score 15.4; DB 15; Length 676;  
Best Local Similarity 94.1%; Pred. No. 8.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGGAAGGATAAATA 17  
|||||  
DB 233 TGTGGAAGGATAAATA 217

Search completed: March 25, 2004, 15:52:49  
Job time : 15.7262 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 07:40:23 ; Search time 1023.24 Seconds  
(without alignments)  
4093.601 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1250\_2235  
Perfect score: 986  
Sequence: 1 ctcgcatccatccagcgc.....gagccgtctcggaacagcaga 986

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003as:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	986	100.0	3289	4	Aah26570 Human win
2	986	100.0	6458	6	Abk48984 Genomic D
3	290	29.4	6021	6	Abk48986 Genomic D
4	125	12.7	125	7	Aad48767 Human FOX
5	79.8	8.1	2000	7	Ada71938 Rice gene
6	75	7.6	75	7	Aad48768 Human FOX
7	74.8	7.6	114955	2	Aax53491 Human ade
8	73	7.4	1117	9	Adc86688 Human GPC
9	70	7.1	687	6	Abq19226 Oligonucle
10	70	7.1	687	6	Abq19227 Oligonucle
11	69.4	7.0	114955	2	Aax53491 Human ade
12	69.2	7.0	3183	9	Aad87060 Human GPC
13	69	7.0	1000	3	Aaa02484 Human col
14	68.4	6.9	712	6	Abq40859 Oligonucle
15	68.4	6.9	712	6	Abq40858 Oligonucle
16	67.6	6.9	1172	6	Abq31995 Oligonucle
17	67.6	6.9	1172	6	Abq31994 Oligonucle
18	67.4	6.8	1416	7	Abz20967 Animal te
19	67.2	6.8	12733	6	Abk98631 Vector PE
20	67.2	6.8	12733	8	Accl3882 L. lactis
21	67.2	6.8	12739	6	Abk98592 Vector pB
22	67.2	6.8	12739	8	Accl3843 Plasmid p
23	67	6.8	1172	6	Abq31995 Oligonucle

24	67	6.8	1172	6	ABQ31996	Abq31996 Oligonucle
25	67	6.8	1172	6	ABQ31996	Abq31996 Oligonucle
26	67	6.8	1172	6	ABQ31997	Abq31997 Oligonucle
27	67	6.8	1172	6	ABQ31997	Abq31997 Oligonucle
28	67	6.8	1172	6	ABQ31994	Abq31994 Oligonucle
29	65.2	6.6	72332	8	ADA02552	Ada02552 Human WNT
30	65.2	6.6	72332	9	ADB72290	Adb72290 Human WNT
31	64.2	6.5	600	6	ABQ52497	Abq52497 Oligonucle
32	64.2	6.5	600	6	ABQ52496	Abq52496 Oligonucle
33	64.2	6.5	114793	4	AAD08215	Aad08215 Human gen
34	64	6.5	11849	4	AAK82792	Aak82792 Human imm
35	63.8	6.5	615	7	ACA23975	Aca23975 Prokaryot
36	63.8	6.5	4545	9	ADE52577	Ade52577 Human SQV
37	63.6	6.5	1951	6	ABQ19255	Abq19255 Oligonucle
38	63.6	6.5	1951	6	ABQ19254	Abq19254 Oligonucle
39	63.6	6.5	5452	9	ADC86736	Adc86736 Human GPC
40	63.4	6.4	94720	8	ADA02654	Ada02654 Human STA
41	63.4	6.4	94720	9	ADB72392	Adb72392 Human STA
42	63.2	6.4	4523	9	ADE60227	Ade60227 Human gen
43	63.2	6.4	4523	9	ADE60229	Ade60229 Human gen
44	63.2	6.4	4523	9	ADE60231	Ade60231 Human gen
45	63.2	6.4	4523	9	ADE60233	Ade60233 Human gen

## ALIGNMENTS

## RESULT 1

AAH26570  
ID AAH26570 standard; cDNA; 3289 BP.

XX  
AC AAH26570;

DT 12-NOV-2001 (first entry)

DE Human winged helix protein FOXC2 cDNA.

XX  
KW FOXC2; human; winged helix protein; FKHL14; transcription factor; forkhead gene; energy balance; adipose tissue; transgenic animal; obesity; malnutrition; hyperlipidaemia; diabetes; hypertension; antidiabetic; anorectic; hypolipaeamic; hypotensive; gene therapy; ss.

OS Homo sapiens.

FH Key Location/Qualifiers  
FT CDS 1197..12705  
FT /\*tag= a

XX  
PN WO200160853-Al.

XX  
PD 23-AUG-2001.

XX  
PF 16-FEB-2001; 2001WO-SE000339.

XX  
PR 18-FEB-2000; 2000SE-00000531.

XX  
PR 26-MAY-2000; 2000SE-00001982.

XX  
PR 06-JUN-2000; 2000US-00587945.

XX  
PR 14-DEC-2000; 2000SE-00004629.

XX  
(PRAA ) PHARMACIA AB.

XX  
PI Enerbaeck S, Carlsson P;

XX  
DR WPI; 2001-557641/62.

XX  
DR P-PSDB; AAB82851.

XX  
PT New construct comprising a human FOXC2 gene, useful in gene therapy for treating obesity- or malnutrition-related diseases (e.g. obesity or hyperlipidemia), as well as for identifying compounds useful in treating these diseases.

XX  
PS Claim 4(a); Page 82-85; 92pp; English.

CC The present sequence is that of cDNA encoding human FOXC2 winged helix  
CC protein (see A882851). The cDNA was isolated from a human fat cell cDNA  
CC library using a mix of cDNA probes corresponding to DNA-binding domains  
CC from different winged helix proteins. FOXC2 (also designated FKHL14) was  
CC one of the genes identified. Northern blotting showed FOXC2 to be  
CC expressed exclusively in human adult adipose tissue. A major role for  
CC FOXC2 in regulating energy balance and adiposity was demonstrated. The  
CC invention provides transgenic non-human mammals that are capable of  
CC expressing the human FOXC2 gene in their adipose tissue. Methods are  
CC provided for identifying compounds useful for the treatment of medical  
CC conditions related to obesity, such as obesity, non-insulin dependent  
CC diabetes, hypertension and hyperlipidaemia (claimed). Such compounds are  
CC stimulate the biological activity of a human FOXC2 polypeptide. Methods  
CC are also provided for identifying compounds useful for the treatment of  
CC medical conditions related to malnutrition, such as anorexia (claimed).  
CC Such compounds decrease human FOXC2 gene expression or FOXC2 protein  
CC activity. Gene therapy may be used as a method of treatment  
XX  
SQ Sequence 3289 BP; 639 A; 1125 C; 925 G; 600 T; 0 U; 0 Other;

Query Match 100.0%; Score 986; DB 4; Length 3289;  
Best Local Similarity 100.0%; Pred. No. 1.6e-202;  
Matches 986; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCATTCCAAATCCAGCGGTTTGCTTTTGAATCCATTACACCTGGGCCCCCAATAATTA 60  
DB 212 CTGCCATTCCAAATCCAGCGGTTTGCTTTTGAATCCATTACACCTGGGCCCCCAATAATTA 271  
QY 61 GGNAAATCTAATTTTCGCTTCATCCTCACTCAATTAAGAAATATGCCAGGATCAATGCT 120  
DB 272 GGAATATCTAATTTTCGCTTCATCCTCACTCAATTAAGAAATATGCCAGGATCAATGCT 331  
QY 121 ACTTCAAGAGTCTTTGGGAGAGATATTTACTTATTAATCAATCTATTTATATTTTCA 180  
DB 332 ACTTCAAGAGTCTTTGGGAGAGATATTTACTTATTAATCAATCTATTTATATTTCA 391  
QY 181 AATGTATTTTAAACAGAGGAAATGCGTATCTTTTGTGGGATGTGGGCCCAT 240  
DB 392 AATGTATTTTAAACAGAGGAAATGCGTATCTTTTGTGGGATGTGGGCCCAT 451  
QY 241 TCACCAAAATGTGATCATATAAATAAATTTAATAAGATATACTTTTAAAAAGTTTCA 300  
DB 452 TCACCAAAATGTGATCATATAAATAAATTTAATAAGATATACTTTTAAAAAGTTTCA 511  
QY 301 AGTGAACAGGAGTCCGCGGAGGCGCGGGGGGGGCTTTAGAGCGACGGAATTCCT 360  
DB 512 AGTGAACAGGAGTCCGCGGAGGCGCGGGGGGGGCTTTAGAGCGACGGAATTCCT 571  
QY 361 GCGCTCTCGCCCGGATTCGCGCGGACCTCTCTCAGCTGCGGGGTGATGGCTCAAAAGT 420  
DB 572 GCGCTCTCGCCCGGATTCGCGCGGACCTCTCTCAGCTGCGGGGTGATGGCTCAAAAGT 631  
QY 421 TCGGAGGGGGGGTGGCGCGAGGAAAGTAAAACTCGCTTTTTCAGCAAGACACTTTTGA 480  
DB 632 TCGGAGGGGGGGTGGCGCGAGGAAAGTAAAACTCGCTTTTTCAGCAAGACACTTTTGA 691  
QY 481 AACTTTTCCCAATCCCTAAAGGACTTGGCTCTTTTCTGGGCTCAGCGGGGAGCGG 540  
DB 692 AACTTTTCCCAATCCCTAAAGGACTTGGCTCTTTTCTGGGCTCAGCGGGGAGCGG 751  
QY 541 CTCGGACCCCGGCGCTGACCTCGGGGCTGCGGATTCGCTGGGGGCTTGAGAGCCCTC 600  
DB 752 CTCGGACCCCGGCGCTGACCTCGGGGCTGCGGATTCGCTGGGGGCTTGAGAGCCCTC 811  
QY 601 CTGCGCCCTCTCTGCGCGGGCGGAGGCTCACTTGTGTCCTCCAGCGCGGCTCTCCG 660  
DB 812 CTGCGCCCTCTCTGCGCGGGCGGAGGCTCACTTGTGTCCTCCAGCGCGGCTCTCCG 871  
QY 661 CTGGGTCCGCGCGGCGGCTGCGCGGCTGCGCGGCGGCTGCTGAGAGCCAGCAGG 720  
DB 872 CTGGGTCCGCGCGGCGGCTGCGCGGCTGCGCGGCGGCTGCTGAGAGCCAGCAGG 931  
QY 721 AGCGGGGCGGCGCTGCGCTTGGCGGGGCGGCGGCTTCAGGATGCCGATCCGCCCGGTC 780

DB 932 AGCGGGGCGGCGCTGCGCTTTGCCCGGGGCGCCCTTCAGGATGCCGATCCCGCGGTC 991  
QY 781 CGCTGAAAGCGGCGGCGGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
DB 992 CGCTGAAAGCGGCGGCGGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1051  
QY 841 AGGCTGCAGAGAGACCGGCGGCGGCGGCTGCTCGGCTGCGGCTGCTGCTGCTGCTGCT 900  
DB 1052 AGGCTGCAGAGAGACCGGCGGCGGCGGCTGCTCGGCTGCGGCTGCTGCTGCTGCTGCT 1111  
QY 901 CGGCTCTCTCGCTCTCAGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 1112 CGGCTCTCTCGCTCTCAGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171  
QY 961 GCCGCGGAGCGGCTTCGGAAGCAGCA 986  
DB 1172 GCCGCGGAGCGGCTTCGGAAGCAGCA 1197  
RESULT 2  
ABX48984  
ID ABX48984 standard; DNA; 6458 BP.  
XX  
AC ABX48984;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Genomic DNA encoding human transcription factor FOXC2.  
XX  
KW Transcription factor; FOXC2; antidiabetic; anorectic; antilipemic;  
KW cardiovascular; FOXC2 polypeptide expression modulator; FOXC2 promoter;  
KW diabetes mellitus; obesity; hypercholesterolaemia; dyslipidaemia;  
KW cardiovascular disease; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
exon 1. .215  
FT /tag= b  
FT /note= "First exon according to the alternative  
FT transcript"  
FT 1. .186  
FT /tag= a  
FT misc\_feature 187. .215  
FT /tag= c  
FT /note= "Region coding for 5' part of alternative protein"  
FT 215. .216  
FT misc\_signal /tag= d  
FT /note= "Alternative first exon splice site"  
FT misc\_difference 216. .475  
FT /tag= e  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in Claim 16"  
FT misc\_difference 223. .231  
FT /tag= f  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in Claim 12"  
FT misc\_difference 359. .375  
FT /tag= g  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in Claim 13"  
FT misc\_difference 378. .402  
FT /tag= h  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in Claim 14"  
FT misc\_difference 403. .423  
FT /tag= i  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in Claim 15"  
FT misc\_difference 1250. .2235  
FT /tag= k  
FT /note= "Fragment of the FOXC2 promoter. Specifically







PT Identifying RNA-binding molecule by predicting structure of RNA fragment,  
PT synthesizing DNA fragment corresponding to predicted RNA structure,  
PT performing reporter gene assay after placing the DNA upstream of reporter  
gene.

XX Claim 12; Page 29; 35pp; English.

XX The present invention relates to a method of identifying RNA-binding  
CC molecule comprising predicting the structure of RNA-fragment, selecting  
CC suitable predicted RNA-fragment with an individual stem, synthesizing a  
CC DNA-fragment corresponding to the RNA fragment, inserting the DNA  
CC fragment in upstream proximity of reporter assay gene to form reporter  
CC construct and performing a reporter gene assay which detects interaction  
CC between a molecule to be tested for RNA-binding and RNA fragment of the  
CC reporter construct. The method is useful for identifying an RNA binding  
CC molecule. The present sequence is human FOXC2 RNA fragment. This sequence  
CC is used to illustrate the method of the invention

XX SQ Sequence 125 BP; 7 A; 65 C; 32 G; 0 T; 21 U; 0 Other;

Query Match 12.7%; Score 125; DB 7; Length 125;

Best Local Similarity 83.2%; Pred. No. 1.6e-17;

Matches 104; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

Qy 861 CCGCCCTCCGCTCCCTCTCTCCCTCTGCTCTCTGCGCTCTCTGCTCTCAGG 920

Db 1 CCGCCCTCCGCTCCCTCTCTCCCTCTGCTCTCTGCGCTCTCTGCTCTCAGG 60

Qy 921 GCCCCTCCGCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGAA 980

Db 61 GCCCCTCCGCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGAA 120

Qy 981 GCAGC 985

Db 121 GCAGC 125

#### RESULT 5

ADA71938/c

ID ADA71938 standard; DNA; 2000 BP.

XX AC ADA71938;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.

XX OS Oryza sativa.

XX PN WO200300898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

XX Claim 27; SEQ ID NO 5263; 899pp; English.

XX

CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 8.1%; Score 79.8; DB 7; Length 2000;

Best Local Similarity 10.0%; Pred. No. 1.9e-07;

Matches 98; Conservative 460; Mismatches 41; Indels 7; Gaps 4;

Qy 6 ATTCCAATCCAGCGGTTTGGCTTTTGAATCCATTACACCTGGGCCCCCATATAGGAAA 65

Db 985 WYWAYKRAYAWSRSRKTWCTGGSRMATYCGTKWAAAGRWWRMAWCWCMKWKVT 926

Qy 66 TCTAATTATTCGCTTCATCCTCATTAATAAGAAAATGTCAGGATCATGTACTTA 125

Db 925 SCWKKYRTWSCVY--TMMGAMRYAYAMRRRTYKNSWRMYWTMTKWTMTMTCM 868

Qy 126 CAAGGTCTTTGGGAGAGATATTTACTCTATTATCCATTCTATTATATTTCAAAATG 185

Db 867 CNAKYMATGATWMMRYTMYTYCANTCAKCYKAMTKMTTACAWRATSWRPM 808

Qy 186 ATTTTATTAAGAGAAAGTGCGTATCTTTTGTGGCATGTGGGCCATC--- 242

Db 807 AGMRKRYKMKRAYWWRWRCWKAGARMKSRVWKKYATRYWQWAMTWMSWR 748

Qy 243 ACCAAAATGTGATCATAAAATAAATTTTAAAGATATACTTTTAAAGATTTCAG 302

Db 747 WKSRYNSGMRWSAWRYCSRMAKTKYASSARWTKRAKSTRYRRRYRRKRGTY 688

Qy 303 TGAAGACGAGTCGCCGCGAGCGCGGCGCGGCTTTAGAGCCGAGGATTCCTGC 362

Db 687 RYRYNRSRMTARMSKRRKAGASMKSCWYWRGASMYKSKYSCSAKCKTRYMTSS 628

Qy 363 GCTCTCCGCGGATTTGGCGCGGCTCTCTCAGCTCCGCGGTGATGGCTCAAAGTTC 422

Db 627 YNSTGMYGMYSSYSMSWTSKSYNGKMTCTNYTSMKSTERSKMGSRWSGMSRMYRWK 568

Qy 423 CGGAGCGGCGGTGGCCGAGGAAAGTAAACCTCGCTTTCAGCAAGAGACTTTTGA 482

Db 567 XMRKRYMYMKWCTWRCWCYRWGYTMYTTSRSMYTGRYKARYTSKRYMYKYRK 508

Qy 483 CTTTCCCAATCCCTAAAGGAGCTTGGCTCTTTTCTGGGCTCAGCGGCGCGCT 542

Db 507 YCWYYTGYMKKCSYMYRYGYCKACKCCYANCAKAAYSGNMYWYRYKSKWMSMKY 448

Qy 543 CGGACCCCGCGCGCTGACCTCGGGGCTCGCGATTC-GCTGGGGGCTTGAGAGCTCC 601

Db 447 WSMYKCRSMKYGAKCYGCMWYCSYGKMYTYMSYKYSRYKRYMYWYKMY 388

Qy 602 TCGCGCCCTCTCGCGCGCGGCGGAGGTTCACCTTGGTCCCGCGCGCGCTCCG 661

Db 387 WMYTSAVSSMMTWYYAKYWKYKRGTMWYKSKYKYYCTWCMCMCRCYRWLKM 328

Qy 662 TGGGTCCGCGCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGGAGCAGG 721

Db 327 MKRTKYSKRYATCYWCCCYRKGYWYRRSMRMTAGKMRWSRWCYSRYWY 268

Qy 722 GCGGCGCGGCTCGCTTCCCGGCGCGCTCCAGGATCGCGCGCGCTCC 781

Db 267 YKOWMKSYMSYWARSSGTWMSRAKRTYKGYSTSRRAKMRACRMYSACRYSTSY 208

Qy 782 CTTGAACCGCGCGCTCGCTCGCGCGGAGCAGGACCGCGCGCTCGCGCGG 841

Db 207 YC-GCSYCGSSKWKYMSKSMRMTCSWCSCTCYTGAGCWCMSMWTMGSCGCT 149





```
PR 05-SEP-2000; 2000DE-01044543.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one member,
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
XX degree of hybridisation to both classes is determined from the label on
XX the amplicon. From the ratio of labels hybridised to the two classes of
XX oligomers, the degree of methylation is calculated. The method is used:
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
XX and of a wide range of diseases, e.g. cancer, disorders of the central
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
XX particularly by detecting mutations or single nucleotide polymorphisms
XX (SNP's); and (ii) for differentiation of cell or tissue types and for
XX investigating cell differentiation. The method allows the methylation
XX status of many C residues to be determined simultaneously. ABQ13410-
XX ABQ54121 represent genomic DNA sequences used to illustrate the method
XX for determining the degree of cytosine methylation described in the
XX disclosure of the invention
XX
XX Sequence 687 BP; 11 A; 123 C; 446 G; 88 T; 0 U; 19 Other;
XX
XX Query Match 7.1%; Score 70; DB 6; Length 687;
XX Best Local Similarity 48.7%; Pred. No. 1.9e-05;
XX Matches 175; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
XX
QY 597 CTTCTGCGCCCTCTCTCGCGCGCGGCGAGGGTCCACCTTGTCTCCAGCGCGCGCGTC 656
Db 514 CCCCCCGCGCCCGAGCGCCCGCCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 455
QY 657 TCCGCTGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 716
Db 454 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 395
QY 717 GAGGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 776
Db 394 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
QY 777 GGTTCGCTGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 836
Db 334 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
QY 837 CCGGAGGCTCCAGGAGACGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896
Db 274 CGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 215
QY 897 CTCTCGCGTCTCTCGCTCTCAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 955
Db 214 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 156
XX
XX RESULT 10
XX ABQ19227
XX ID ABQ19227 standard; DNA; 687 BP.
XX
XX ABQ19227;
XX
XX 12-JUL-2002 (first entry)
```

QY 837 CCGAGAGCTCCAGGAGACCGGGCGCCCTCCCGCTCCCTCTCCCTCTCCCTCTGGCT 896  
Db 414 CGAACGCCCCCGCCCCCGCCCGACCGAACGACCGCCCCCGCCCCCGCCCCGCGC 473  
QY 897 CTCTGGGCTCTCTCGCTCTCAGGGCCCCCTCGCTCCCCCGCGCGAGTCCGTGCGG 955  
Db 474 CCGCGCGCCCCCGGAAACGACGACCCGCCCCCGAGCGCCCGCCCCCGCGCG 532

## RESULT 11

AA53491/c  
ID AA53491 standard; DNA; 114955 BP.

AC AA53491;

XX 05-JUL-1999 (first entry)

DE Human adenosine A1 receptor antisense oligonucleotide fragment.

XX Antisense oligonucleotide; multiple target; antisense treatment;  
KW impaired respiration; inflammation; lung disease;  
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
KW acute asthma; allergy; asthma; impeded respiration;  
KW respiratory distress syndrome; pain; cystic fibrosis;  
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
KW prostate cancer; ss.

XX Synthetic.

XX WO9913886-A1.

XX 25-MAR-1999.

XX 17-SEP-1998; 98WO-US019419.

XX 17-SEP-1997; 97US-0059160P.

PR 09-JUN-1998; 98US-00093972.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 1998-229400/19.

XX New antisense oligonucleotides used in treatment of, e.g. pulmonary  
PT vasoconstriction.

PS Disclosure; Page 37; 120pp; English.

XX The specification describes antisense oligonucleotides (AA52869-X55271)  
CC directed against at least 2 mRNAs selected from target genes, coding and  
CC non-coding regions of RNAs corresponding to target genes, gene initiation  
CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-  
CC end and the juxta-section between coding and non-coding regions and all  
CC segments of RNAs encoding proteins associated with one or more diseases,  
CC conditions or mixtures. The antisense oligonucleotides may be derived  
CC from sequences AA5272-74. These multiple target oligonucleotides  
CC (specifically AA55180-271) can be used for the antisense treatment of  
CC diseases and conditions. Typical diseases and conditions are those  
CC associated with impaired respiration and inflammation, including lung  
CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,  
CC acute asthma, allergies, asthma, impeded respiration, respiratory  
CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,  
CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary  
CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.  
CC colon cancer, breast cancer, lung cancer, pancreatic cancer,  
CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as  
CC well as all types of cancers which may metastasize or have metastasized  
CC to the lungs, including breast and prostate cancer

XX SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;  
Query Match 7.0%; Score 69.4; DB 2; Length 114955;  
Best Local Similarity 37.0%; Pred. No. 0.0001;  
Matches 166; Conservative 25; Mismatches 258; Indels 0; Gaps 0;  
QY 523 GGCTCAGGGGGGAGCGCTCGGACCGCGCGCGCTGACCTCGGGGTGCGATTGCT 582  
Db 105661 GGGCAAGCGGGCGCGGGGCGGAGCCAGGGGCCNNHNNSCGGGCGGGCGGCAAGC 105602  
QY 583 GGGGCTTTGAGAGAGCTCTCTCGCCCTCTCTCGCGCGGGCGAGGGTCCACCTTGGTCCC 642  
Db 105601 CGGGCGCGGGCGGAGCCAGGGGCCNNHNNSCGGGCGGGCGGCAAGCGGGCGC 105542  
QY 643 CAGCGCGGGGCTCTCTCGCTGGTCCGCGCGCGCGCTGCGCGGTGCGCGCGCGG 702  
Db 105541 CGGGCGGAGCCAGGGGCCNNHNNSVGGGAGCGCGGCCNNHNNSCGGGCGGGCC 105482  
QY 703 GTCTTGAGAGCCAGGAGCGGGCGCGCTGCGCTTGGCGGGCGCGCTCCAGG 762  
Db 105481 NNHNNSCGGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105422  
QY 763 ATGCGATCGCGCGGTCTCTGAAAGCGCGCGCGCTCTCGCGCGGAGCGAGCAGAC 822  
Db 105421 NNNSCGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105362  
QY 823 CGCGACCTCTCGCCCGGAGGCTGCGAGAGCGGGCGCGCGCTCCCGCTCCCTCT 882  
Db 105361 GCGNNHNNSCGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105302  
QY 883 CTCCCT 942  
Db 105301 GCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105242  
QY 943 CAGTCTCTGCGGAGGGCGCGCGGAGCC 971  
Db 105241 CCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105213

## RESULT 12

ADC87060/c

ID ADC87060 standard; DNA; 3163 BP.

XX ADC87060;

XX 01-JAN-2004 (first entry)

XX Human GPCR gene SEQ ID NO:1513.

XX ds; gene; human; GPCR;

XX guanosine triphosphate-binding protein coupled receptor; gene therapy.

XX Homo sapiens.

XX EP1270724-A2.

XX 02-JAN-2003.

XX 18-JUN-2002; 2002EP-00013517.

XX 18-JUN-2001; 2001JP-00246789.

XX (NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

XX Suwa M, Asai K, Akiyama Y, Aburatani H;

XX WPI; 2003-315783/31.

XX P-PSDB; ADC87061.

XX New polynucleotide, useful for preparing a composition for treating a  
PT patient in need of increased or suppressed activity or expression of the





XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX SNP; cell differentiation; ds.  
OS Homo sapiens.  
XX WO200218632-A2.  
XX 07-MAR-2002.  
XX 01-SEP-2001; 2001WO-EP010074.  
XX 01-SEP-2000; 2000DE-01043826.  
XX 05-SEP-2000; 2000DE-01044543.  
XX (EPIG-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
XX WPI; 2002-371829/40.  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
XX diagnosis and prognosis, comprises selective hybridization of amplicons  
XX from chemically treated DNA.  
XX Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX This invention describes a novel method for determining the degree of  
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
XX genomic sample of DNA. The sample is treated chemically to convert  
XX cytosine (C) but not methylated C, to uracil, then part of the genomic  
XX DNA that contains the target C is amplified to form a labeled amplicon.  
XX The amplicon is hybridized to two classes, each with at least one member,  
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
XX degree of hybridization to both classes is determined from the label on  
XX the amplicon. From the ratio of labels hybridized to the two classes of  
XX oligomers, the degree of methylation is calculated. The method is used:  
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
XX and of a wide range of diseases, e.g. cancer, disorders of the central  
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
XX particularly by detecting mutations or single nucleotide polymorphisms  
XX (SNP's); and (ii) for differentiation of cell or tissue types and for  
XX investigating cell differentiation. The method allows the methylation  
XX status of many C residues to be determined simultaneously. ABQ13410-  
XX ABQ54121 represent genomic DNA sequences used to illustrate the method  
XX for determining the degree of cytosine methylation described in the  
XX disclosure of the invention  
XX Sequence 712 BP; 47 A; 456 C; 181 G; 12 T; 0 U; 16 Other;

Query Match 6.9%; Score 68.4; DB 6; Length 712;  
Best Local Similarity 48.6%; Pred. NO. 4.1e-05;  
Matches 208; Conservative 0; Mismatches 217; Indels 3; Gaps 1;  
QY 524 GCTCAGCGGGGACCGCTCGACCCCGCGCGGTGACCTCGGGGTGCGATTCGCTG 583  
DB 260 GCTCGCGCCCGCTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 319  
QY 584 GGGGCTTGAGAGGCTCTCGCGCCC---CTCCTCGCGCGCGCGCGCGCGCGCGCG 640  
DB 320 GCGCGCGCTCG 379  
QY 641 CCCAGCGCGCGCGGTCTCGGTGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700  
DB 380 CCCCG 439  
QY 701 GGGTCTCGAGCGAGGAGCGGGCGCGCGGTGCGCTTGGCCCGCGCGCGCGCGCGCG 760  
DB 440 CGAAGCGCGAAGCG 499  
QY 761 GGATGCGGATCGCGCGCGCTCGCTGAAAGCGCGCGCGCGCGCGCGCGCGCGCG 820

DB 500 CCCCCCGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 559  
QY 821 ACCGCGACCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 880  
DB 560 CG 619  
QY 881 CTCTCCCCCTCTGGGTCTCTCGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 940  
DB 620 CCGCGCGACCG 979  
QY 941 CGCAGTCC 948  
DB 680 NCGGACCC 687  
RESULT 15  
ABQ40858/c  
ID ABQ40858 standard; DNA; 712 BP.  
XX ABQ40858;  
XX 12-JUL-2002 (first entry)  
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 27449.  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX SNP; cell differentiation; ds.  
XX Homo sapiens.  
XX WO200218632-A2.  
XX 07-MAR-2002.  
XX 01-SEP-2001; 2001WO-EP010074.  
XX 01-SEP-2000; 2000DE-01043826.  
XX 05-SEP-2000; 2000DE-01044543.  
XX (EPIG-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
XX WPI; 2002-371829/40.  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
XX diagnosis and prognosis, comprises selective hybridization of amplicons  
XX from chemically treated DNA.  
XX Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX This invention describes a novel method for determining the degree of  
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
XX genomic sample of DNA. The sample is treated chemically to convert  
XX cytosine (C) but not methylated C, to uracil, then part of the genomic  
XX DNA that contains the target C is amplified to form a labeled amplicon.  
XX The amplicon is hybridized to two classes, each with at least one member,  
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
XX degree of hybridization to both classes is determined from the label on  
XX the amplicon. From the ratio of labels hybridized to the two classes of  
XX oligomers, the degree of methylation is calculated. The method is used:  
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
XX and of a wide range of diseases, e.g. cancer, disorders of the central  
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
XX particularly by detecting mutations or single nucleotide polymorphisms  
XX (SNP's); and (ii) for differentiation of cell or tissue types and for  
XX investigating cell differentiation. The method allows the methylation  
XX status of many C residues to be determined simultaneously. ABQ13410-  
XX ABQ54121 represent genomic DNA sequences used to illustrate the method  
XX for determining the degree of cytosine methylation described in the



CC disclosure of the invention  
 XX  
 SQ Sequence 712 BP; 12 A; 181 C; 456 G; 47 T; 0 U; 16 Other;  
 Query Match 6.9%; Score 68.4; DB 6; Length 712;  
 Best Local Similarity 48.6%; Pred. No. 4.1e-05;  
 Matches 208; Conservative 0; Mismatches 217; Indels 3; Gaps 1;  
 Qy 524 GCTCAGGGGGGACGCGCTCGGACCGCGCGCGCTGACCTCGGGGCTGCGGATTGCTG 583  
 Db |||||  
 Qy 453 GCTCCGCGCGCGCTCG 394  
 Db |||||  
 Qy 584 GGGCTTGGAGAGCTCTCTGCGCC---CTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640  
 Db |||||  
 Qy 393 GCGGCGGCTCCG 334  
 Db |||||  
 Qy 641 CCGAGGCGCGCGCGCTCTCGGCTGGGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700  
 Db |||||  
 Qy 333 CCG 274  
 Db |||||  
 Qy 701 GGGTCTCTGAGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 760  
 Db |||||  
 Qy 273 CGRACGCGGACCGAN 214  
 Db |||||  
 Qy 761 GGATGCGGATCG 820  
 Db |||||  
 Qy 213 CCG 154  
 Db |||||  
 Qy 821 ACCGCGCACCTCGCTC 880  
 Db |||||  
 Qy 153 CG 94  
 Db |||||  
 Qy 881 CTCTCCCG 940  
 Db |||||  
 Qy 93 CCGCGCGGACCGCGCGGACCGCGCGGACCGCGCGGACCGCGCGGACCGCGCGGACCGCG 34  
 Db |||||  
 Qy 941 CGCAGTCC 948  
 Db |||||  
 Qy 33 NCGAGCC 26

Search completed: March 25, 2004, 10:25:05  
 Job time : 1027.24 secs



```
QY 945 GTCCGTCTCGAGGGCGCCCGGAGCGCTCTCGG 978
Db 22179 CCCCCCCCCCGGCCCCCCCCCGCGGGCGG 22212

RESULT 2
US-09-616-289-48
; Sequence 48, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Ann M.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-09-616-289-48

Query Match 6.0%; Score 59; DB 4; Length 2561;
Best Local Similarity 52.2%; Pred. No. 0.00013;
Matches 180; Conservative 0; Mismatches 160; Indels 5; Gaps 2;

QY 605 GCCCTCTCTCGCGGGCGGAGGGTCCACCTTGGTCCCGAGGCGCGGCTCTCGCTGG 664
Db 576 GCCCCCGCGCGCGCCCGCCCGCGGGGGCCCGCGCGCGCGCGCGCGCGCGCGCGCC 635
QY 665 GTCCGCGCGCGCGCGCTGCTGCGCGCGCGTCCGCGCGCGTCTCGAGGAGC 723
Db 636 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 695
QY 724 GGGGCGCGCGCTCGCTTGGTCCCGGGGGCGGCGCTCCAGGATG----CCGATCGCGCGGT 779
Db 696 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 755
QY 780 CCCTGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
Db 756 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 815
QY 840 GAGGCTGCCAGGAGACGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
Db 816 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 875
QY 900 TCGCGCTCTCTCGCTCTCAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 944
Db 876 CGGAGTTCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 920

RESULT 3
US-09-791-211-10
; Sequence 10, Application US/09791211
; Patent No. 6448080
```

```
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 98844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 24962
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 64383
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65468
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65471
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 87130
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 89049
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
US-09-791-211-10

Query Match 6.0%; Score 58.8; DB 4; Length 98844;
Best Local Similarity 52.9%; Pred. No. 0.00059;
Matches 126; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 729 CGGCGCTGCGCTTCCCGGGCGCGCGCTCCAGGATCCGATCGCGCGCTCCGCTGAAA 789
Db 13404 CCGGCGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13463
QY 789 GCGCGCGCGCGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 848
Db 13464 GCGCGCGCTCGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13523
QY 849 AGGAGACGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 908
Db 13524 CGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 13583
QY 909 CTCGCTCTCAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 966
Db 13584 CACATCATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13641

RESULT 4
US-09-165-264-13/c
; Sequence 13, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiyah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
```





```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-690-473-1

Query Match      5.6%; Score 55; DB 2; Length 4257;
Best Local Similarity 46.5%; Pred. No. 0.0014;
Matches 178; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 600 CTTGCGCCCTCTCTCGCGCGCGGAGGTTCCACCTTGGTCCCGAGCGCGCGCTCTCC 659
DB 2364 CTTGGGCCCCCGCTGCGCGGGACCCCGCGCTTCCGAGATCCCGCGCGCGCGCGC 2423
QY 660 GCTGGGTCCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 719
DB 2424 GGACCTGCTGTTTGACAAACAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCG 2483
QY 720 GAGCGGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 779
DB 2484 GGACGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2543
QY 780 CGCTGAAAGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
DB 2544 GCGCAAGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2603
QY 840 GAGCGTCCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
DB 2604 GAAGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2663
QY 900 TCGCGCTCTCTCGCTCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
DB 2664 GCGCCCTCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2723
QY 960 CGCGCGCGCGCGCTCTCGGAAGC 982
DB 2724 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2746

RESULT 10
US-09-259-821A-1
; Sequence 1, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259, 821A
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
```

```
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
; US-09-259-821A-1

Query Match      5.6%; Score 55; DB 3; Length 4257;
Best Local Similarity 46.5%; Pred. No. 0.0014;
Matches 178; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 600 CTTGCGCCCTCTCTCGCGCGCGGAGGTTCCACCTTGGTCCCGAGCGCGCGCTCTCC 659
DB 2364 CTTGGGCCCCCGCTGCGCGGGACCCCGCGCTTCCGAGATCCCGCGCGCGCGCGC 2423
QY 660 GCTGGGTCCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 719
DB 2424 GGACCTGCTGTTTGACAAACAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCG 2483
QY 720 GAGCGGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 779
DB 2484 GGACGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2543
QY 780 CGCTGAAAGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
DB 2544 GCGCAAGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2603
QY 840 GAGCGTCCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
DB 2604 GAAGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2663
QY 900 TCGCGCTCTCTCGCTCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
DB 2664 GCGCCCTCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2723
QY 960 CGCGCGCGCGCGCTCTCGGAAGC 982
DB 2724 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2746

RESULT 11
US-08-843-659-1
; Sequence 1, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARS:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-659-1

Query Match          5.6%; Score 55; DB 3; Length 4257;
Best Local Similarity 46.5%; Pred. No. 0.0014;
Matches 178; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 600 CTTGGGCGCCCTCTCGCGGGGCGAGGGTCCACCTTGTGTCCTCCAGGCGCGGGGTCTCC 659
Db 2364 CTTGGGCGCCGCTGTCGGGGGAGACCGCGCTGCTGAGCTCGCGGGCGCGCGCGCGC 2423
QY 660 GCTGGGTCGCGGGCGCGCCCTGCTCGCGCGCTGCTCGCGCGCTGCTCGCGCGCGCGC 719
Db 2424 GGACCTGCTGTTTGACAAACAGAGAGCTGCGCGCGCTGCTGCGCGCGCGCGCGC 2483
QY 720 GAGCGGGCGCGCGCTGCTGCTGCGGGGCGCGCGCTGCTGCGCGCGCGCGCGCGCGT 779
Db 2484 GGACCGCGCGAGCGCTGCGCGCGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGC 2543
QY 780 CGCTGAAGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 839
Db 2544 GCGCAAGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2603
QY 840 GAGCTGCCAGAGACCGCGGGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGC 899
Db 2604 GAAGAAGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2663
QY 900 TCGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Db 2664 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2723
QY 960 CGCGCGCGCGCGCGCTGCTCGGAAGC 982
Db 2724 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC

RESULT 12
US-08-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; TITLE OF INVENTION: Infections
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339ris
; STREET: One Liberty Place, 48th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,568A
; FILING DATE: 02-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/065,146
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFCI-0029
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; US-08-458-568A-11

Query Match          5.6%; Score 55; DB 1; Length 12001;
Best Local Similarity 46.5%; Pred. No. 0.0021;
Matches 178; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 600 CTTGGGCGCCCTCTCGCGGGGCGAGGGTCCACCTTGTGTCCTCCAGGCGCGGGGTCTCC 659
Db 4310 CTTGGGCGCCGCTGCTGCGGGGAGACCGCGCTGCTGAGCTCCGCGCGCGCGCGCGC 4251
QY 660 GCTGGGTCGCGGGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGT 719
Db 4250 GACCTGCTGTTTGACAAACAGAGAGCTGCGCGCGCTGCTGCGCGCGCGCGCGC 4191
QY 720 GAGCGGGCGCGCGCTGCTGCTGCGGGGCGCGCGCTGCTGCGCGCGCGCGCGCGCGT 779
Db 4190 GAGCGCGCGCGCGCGCTGCTGCGGGGCGCGCGCTGCTGCGCGCGCGCGCGCGCGC 4131
QY 780 CGCTGAAGCGCGCGCGCTGCTCGGAAGC 982
Db 4130 GCGCAAGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 839
QY 840 GAGCTGCCAGAGACCGCGGGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGC 899
Db 4070 GAAGAAGAGCGCGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGC 4011
QY 900 TCGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Db 4010 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 3951
QY 960 CGCGCGCGCGCGCGCTGCTCGGAAGC 982
Db 3950 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC

RESULT 13
US-09-165-264-11/c
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuralayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
; US-09-165-264-11

Query Match          5.5%; Score 54.6; DB 3; Length 320;
Best Local Similarity 48.5%; Pred. No. 0.00062;
Matches 150; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
```

Qy 629 TCCACCTGGTCCAGCGCGGCTCTCGGTGGGTCCGCGCGCGCGCTGCCGC 688  
Db |||  
311 TCAAAATGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 252  
Qy 689 GCTCCGCGCGCGGCTCTGAGAGCAGGAGCGGGCGCGCGCTGCGCTTCCCGGG 748  
Db |||  
251 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 192  
Qy 749 GCGGCGCTCCAGATCCGATCCGCGCGGTCTGCTGAAAGCGGCGCGCTCGGCC 808  
Db |||  
191 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 132  
Qy 809 CGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 868  
Db |||  
131 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 72  
Qy 869 CCGCTCCGCT 928  
Db |||  
71 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 12  
Qy 929 CGTCTCCGCC 937  
Db 11 CCCCCCGCC 3

## RESULT 14

US-09-165-264-12/c  
; Sequence 12, Application US/09165264  
; Patent No. 6197510  
; GENERAL INFORMATION:  
; APPLICANT: Vinayagancorthy, Thuraiyiah  
; TITLE OF INVENTION: Multi-Loci Genomic Analysis  
; FILE REFERENCE: 44747  
; CURRENT APPLICATION NUMBER: US/09/165,264  
; CURRENT FILING DATE: 1998-10-01  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 318  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence  
US-09-165-264-12

Query Match 5.5%; Score 54.4; DB 3; Length 318;  
Best Local Similarity 48.7%; Pred. No. 0.00069;  
Matches 148; Conservative 0; Mismatches 156; Indels 0; Gaps 0;  
Qy 640 CCCAGCGCGCGGCTCTCGGTGGGTCCGCGCGCGCGCGCTGCCGC 699  
Db |||  
306 CACCAACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 247  
Qy 700 CGGTCTCTGAGCAGCAGGAGCGGGCGCGCGCTGCGTTCGCGGCGCGCTCC 759  
Db |||  
246 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 187  
Qy 760 AGGATCGGATCCGCGGTCTGAAAGCGCGCGCGCTGCTCGGCGCGAGCAGC 819  
Db |||  
186 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 127  
Qy 820 GACCGCGACCTCGCGCCCGAGAGCTGCGAGAGACCGGGCGCGCGCTCCCGCT 879  
Db |||  
126 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 67  
Qy 880 CTTCTCCGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 939  
Db |||  
66 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 7  
Qy 940 CCGC 943  
Db 6 CCCC 3

RESULT 15  
US-08-335-865J-7/c  
; Sequence 7, Application US/08335865J  
; Patent No. 6107472  
; GENERAL INFORMATION:  
; APPLICANT: Stacker, Steven A.; Hovens, Christopher M.,  
; APPLICANT: Wilks, Andrew F.  
; TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 666 Fifth Ave  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: ASCII/Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/335,865J  
; FILING DATE: 19-January-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU93/00210  
; FILING DATE: 10-May-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PL2358  
; FILING DATE: 11-May-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6107472man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD-5277  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 318-3100  
; TELEFAX: (212) 752-5958  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3069  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-335-865J-7

Query Match 5.5%; Score 54.4; DB 3; Length 3069;  
Best Local Similarity 55.7%; Pred. No. 0.0017;  
Matches 167; Conservative 0; Mismatches 126; Indels 7; Gaps 3;  
Qy 595 AGCTCTCTGCGCCCTCTCTCGCGCGCGCGCGGTTCACCTTGTCTCCAGCGCGCG-- 652  
Db |||  
350 AGCGCGCGCACCT 291  
Qy 653 CGTCTCTGCGGTCTCGCGCGCGCGCGCTGCGCGCGCGCGCGCGGTCTCTCTCT 708  
Db |||  
290 TCACACTCGGGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231  
Qy 709 GAGCCAGCAGAGAGCGGGCGCGCGCTGCGTTCGCGCGCGCGCGCGCGCGCGCG 768  
Db |||  
230 CAACAGCAAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 171  
Qy 769 ATCCGCGCGGTCTCGTGAAGCGCG-CGCGCTCTGCTCGCGCGCGCGCGCGCG 827  
Db |||  
170 CTCG 111  
Qy 828 ACCCTCGCGCGAGGCTGCGAGAGACCGGGCGCGCGCTCTCTCTCTCTCTCT 887  
Db |||  
110 CGAAGAGAGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 51

Search completed: March 25, 2004, 15:34:17



Job time : 200.942 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 854.12 Seconds  
(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1250\_2235  
Perfect score: 986  
Sequence: 1 cgcgcattccatccagcgc.....gagccgtctcggaagcagca 986

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	986	100.0	3289	9	US-09-963-285-8
2	986	100.0	6458	9	US-09-963-285-1
3	290	29.4	6021	9	US-09-963-285-5
4	125	12.7	125	14	US-10-116-265-15
5	75	7.6	75	14	US-10-116-265-16
6	74.2	7.5	805	12	US-10-424-599-61903
7	73.6	7.5	51657	15	US-10-057-475B-10475
8	73.6	7.5	51657	15	US-10-154-884B-10475
9	73	7.4	1117	14	US-10-017-161-1403
10	73	7.4	1117	15	US-10-292-798-1141
11	69.2	7.0	3163	15	US-10-017-161-1857
12	69.2	7.0	3163	15	US-10-292-798-1513
13	67.4	6.8	771	12	US-10-424-599-107958
14	67.2	6.8	12733	14	US-10-032-393-47
15	67.2	6.8	12733	14	US-10-032-393-8

C 16	67	6.8	721	12	US-10-424-599-139811	Sequence 139811,
C 17	64.8	6.6	51657	15	US-10-057-475B-10475	Sequence 10475, A
C 18	64.8	6.6	51657	15	US-10-154-884B-10475	Sequence 10475, A
C 19	64.4	6.5	497	12	US-10-424-599-133788	Sequence 133788,
C 20	64.2	6.5	114793	14	US-10-148-806-3	Sequence 3, Appli
C 21	63.8	6.5	615	12	US-10-282-122A-11845	Sequence 11845, A
C 22	63.8	6.5	4545	15	US-10-347-470A-29	Sequence 29, Appl
C 23	63.6	6.5	5452	14	US-10-017-161-1481	Sequence 1481, Ap
C 24	63.6	6.5	5452	15	US-10-292-798-1199	Sequence 1189, Ap
C 25	62.6	6.3	815	12	US-10-424-599-20495	Sequence 20495, A
C 26	62.2	6.3	809	12	US-10-424-599-26988	Sequence 26988, A
C 27	62	6.3	606	12	US-10-424-599-95310	Sequence 95310, A
C 28	61.6	6.2	3133	14	US-10-017-161-1483	Sequence 1483, Ap
C 29	61.6	6.2	3133	15	US-10-292-798-1191	Sequence 1191, Ap
C 30	59.8	6.1	152331	13	US-10-095-407-16	Sequence 16, Appl
C 31	59.2	6.0	2307	10	US-09-893-519A-87	Sequence 87, Appl
C 32	59	6.0	2561	9	US-09-976-740-48	Sequence 48, Appl
C 33	59	6.0	2561	12	US-10-671-242-48	Sequence 48, Appl
C 34	59	6.0	2561	13	US-10-023-529-48	Sequence 48, Appl
C 35	59	6.0	2561	13	US-10-023-523-48	Sequence 48, Appl
C 36	59	6.0	2561	15	US-10-616-187-48	Sequence 48, Appl
C 37	59	6.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
C 38	58.8	6.0	580	12	US-10-424-599-2320	Sequence 2320, Ap
C 39	58.8	6.0	2712	9	US-09-963-285-6	Sequence 6, Appli
C 40	57.8	5.9	956	15	US-10-027-832-31508	Sequence 31508, A
C 41	57.6	5.8	778	12	US-10-424-599-54839	Sequence 54839, A
C 42	57	5.8	43058	9	US-09-954-456-292	Sequence 292, App
C 43	57	5.8	43058	9	US-09-954-456-529	Sequence 529, App
C 44	57	5.8	43058	9	US-09-880-107-3950	Sequence 3950, Ap
C 45	57	5.8	155074	13	US-10-026-188-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-963-285-8  
; Sequence 8, Application US/09963285  
; Patent No. US20020090707A1  
; GENERAL INFORMATION:  
; APPLICANT: Enerback, Sven  
; APPLICANT: Krok, Katarina  
; APPLICANT: Rondahl, Lena  
; APPLICANT: Wasserman, Wyeth  
; TITLE OF INVENTION: PROMOTER SEQUENCES  
; FILE REFERENCE: 13425-042001  
; CURRENT APPLICATION NUMBER: US/09/963,285  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: SE 0004102-0  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/238,897  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: SE 0003435-5  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 3289  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-963-285-8

Query Match	100.0%;	Score	986;	DB	9;	Length	3289;
Best Local Similarity	100.0%;	Pred. No.	1.3e-241;				
Matches	986;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	CTGCATTCCAA	TCAGCGGTTGCTTTTGAATCCATTACACCTGGCCCCCAATA	60			
DB	212	CTGCATTCCAA	TCAGCGGTTGCTTTTGAATCCATTACACCTGGCCCCCAATA	271			
QY	61	GGAATCTAATTA	TTCGTTTCATCTAATTAAGAAAAATGTC	120			
DB	272	GGAATCTAATTA	TTCGTTTCATCTAATTAAGAAAAATGTC	331			

```
QY 121 ACTTACAGGCTCTTGGAGAGATATTTTACTCTATTATTAATCAATCTATTTTATATTCA 180
Db 332 ACTTACAGGCTCTTGGAGAGATATTTTACTCTATTATTAATCAATCTATTTTATATTCA 391
QY 181 AATTGATTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGTTGGCATGTTGGCCCAT 240
Db 392 AATTGATTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGTTGGCATGTTGGCCCAT 451
QY 241 TCACCAAAATGTGATCATATAAATAAATTTTAAAGATATAATCTTTTAAAGATTTC 300
Db 452 TCACCAAAATGTGATCATATAAATAAATTTTAAAGATATAATCTTTTAAAGATTTC 511
QY 301 AGTGAAGACGAGTCCGCGGAGGCGGCGGCGGCTCTTAGAGCGACGATTCCT 360
Db 512 AGTGAAGACGAGTCCGCGGAGGCGGCGGCGGCTCTTAGAGCGACGATTCCT 571
QY 361 GCGCTCTCGCCCGATTTGGCGCGGACTCTCTCAGCTGCGGGTGTATGGCTCAAAGT 420
Db 572 GCGCTCTCGCCCGATTTGGCGCGGACTCTCTCAGCTGCGGGTGTATGGCTCAAAGT 631
QY 421 TCCGGAGGGGGTGGCGGAGGAAAGTAAAACTCGCTTTTCAAGAAAGACTTTGA 480
Db 632 TCCGGAGGGGGTGGCGGAGGAAAGTAAAACTCGCTTTTCAAGAAAGACTTTGA 691
QY 481 AACTTTTCCCAATCCCTAAAGGAGCTTGGCCCTCTTTTCTGGGCTCAGCGGGCAGCG 540
Db 692 AACTTTTCCCAATCCCTAAAGGAGCTTGGCCCTCTTTTCTGGGCTCAGCGGGCAGCG 751
QY 541 CTCGGACCCCGGCGCTGACCTCGGGCTGCGGATTCGCTGGGGCTTGAGAGCCCTC 600
Db 752 CTCGGACCCCGGCGCTGACCTCGGGCTGCGGATTCGCTGGGGCTTGAGAGCCCTC 811
QY 601 CTGCGCCCTCTCTCGCGCGGCGAGGCTCCACTTTGGTCCCGCAGCGCGGCTCTCG 660
Db 812 CTGCGCCCTCTCTCGCGCGGCGAGGCTCCACTTTGGTCCCGCAGCGCGGCTCTCG 871
QY 661 CTGGGTCCGGGGCGCGCGCTGCGCGGTGCGCGCGCGGCTCTGAGCGAGGAGG 720
Db 872 CTGGGTCCGGGGCGCGCGCTGCGCGGTGCGCGCGCGGCTCTGAGCGAGGAGG 931
QY 721 AGCGGGCGCGCGCTGCGCTTGGCGGGCGCGCTCCAGGATCGCGATCCCGCGGTC 780
Db 932 AGCGGGCGCGCGCTGCGCTTGGCGGGCGCGCTCCAGGATCGCGATCCCGCGGTC 991
QY 781 CGCTGAAAGCGCGCGCTGCTGCGCGCGAGGAGCGAGCGAGCCCTCGCGCGG 840
Db 992 CGCTGAAAGCGCGCGCTGCTGCGCGCGAGGAGCGAGCGAGCCCTCGCGCGG 1051
QY 841 AGGCTGCCAGGAGACCGGGCGCGCGCTCCCGCTCCCTCTCTCCCGCTCTGCTCT 900
Db 1052 AGGCTGCCAGGAGACCGGGCGCGCGCTCCCGCTCCCTCTCTCCCGCTCTGCTCT 1111
QY 901 CGCGCTCTCTGCTCTCAGGGCGCGCGCTCGCTCCCGCGCGAGTCCGTGCGAGGAG 960
Db 1112 CGCGCTCTCTGCTCTCAGGGCGCGCGCTCGCTCCCGCGCGAGTCCGTGCGAGGAG 1171
QY 961 GCGCGCGCGCTCTCGGAGGAGCA 986
Db 1172 GCGCGCGCGCTCTCGGAGGAGCA 1197
```

## RESULT 2

```
US-09-963-285-1
; Sequence 1: Application US/09963285
; Patent No.: US2002090707A1
; GENERAL INFORMATION:
; APPLICANT: Enerbck, Sven
; APPLICANT: Krook, Katarina
; APPLICANT: Rondahl, Lena
; APPLICANT: Wasserman, Wyeth
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-042001
```

```
; CURRENT APPLICATION NUMBER: US/09/963,285
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: SE 0004102-0
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/238,897
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: SE 0003435-5
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2235) ... (3737)
; US-09-963-285-1
```

Query Match 100.0%; Score 986; DB 9; Length 6458;

Best Local Similarity 100.0%; Pred. No. 1.9e-241;

Matches 986; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTGCCATTCGAATCCAGCGGCTTGTCTTTTGAATCCATTACCTGGGCCCCCATATTA 60
Db 1250 CTGCCATTCGAATCCAGCGGCTTGTCTTTTGAATCCATTACCTGGGCCCCCATATTA 1309
QY 61 GGAATCTAATATTTCCTTCATCCTCATTAATAAGAAAAATGTCCAGAGCATTTGCT 120
Db 1310 GGAATCTAATATTTCCTTCATCCTCATTAATAAGAAAAATGTCCAGAGCATTTGCT 1369
QY 121 ACTTACAGGCTCTTGGAGAGATATTTTACTCTATTATTAATCAATCTATTTTATATTCA 180
Db 1370 ACTTACAGGCTCTTGGAGAGATATTTTACTCTATTATTAATCAATCTATTTTATATTCA 1429
QY 181 AATTGATTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGTTGGCATGTTGGCCCAT 240
Db 1430 AATTGATTTTTTAAACAGAGAAAGTGGCTATCTTTTGTGTTGGCATGTTGGCCCAT 1489
QY 241 TCACCAAAATGTGATCATATAAATAAATTTTAAAGATATAAATCTTTTAAAGATTTC 300
Db 1490 TCACCAAAATGTGATCATATAAATAAATTTTAAAGATATAAATCTTTTAAAGATTTC 1549
QY 301 AGTGAAGACGAGTCCGCGGAGGCGGCGGCGGCTCTTAGAGCGAGCATTCCT 360
Db 1550 AGTGAAGACGAGTCCGCGGAGGCGGCGGCGGCTCTTAGAGCGAGCATTCCT 1609
QY 361 GCGCTCTCTCGCCCGATTTGGCGCGGACTCTCTCAGCTCGCGGCTGATTGGCTCAAAGT 420
Db 1610 GCGCTCTCTCGCCCGATTTGGCGCGGACTCTCTCAGCTCGCGGCTGATTGGCTCAAAGT 1669
QY 421 TCCGGAGGGGGCTGCGCGGAGGAAAGTAAAACTCGCTTTTCAAGAAAGACTTTTGA 480
Db 1670 TCCGGAGGGGGCTGCGCGGAGGAAAGTAAAACTCGCTTTTCAAGAAAGACTTTTGA 1729
QY 481 AACTTTTCCCAATCCCTAAAGGAGCTTGGCTCTTTTCTGCGCTCAGCGGGCAGCG 540
Db 1730 AACTTTTCCCAATCCCTAAAGGAGCTTGGCTCTTTTCTGCGCTCAGCGGGCAGCG 1789
QY 541 CTCGGACCCCGGCGCTGACCTCGGGGTGCGGATTCGCTGGGGCTTTGGAGAGCTC 600
Db 1790 CTCGGACCCCGGCGCTGACCTCGGGGTGCGGATTCGCTGGGGCTTTGGAGAGCTC 1849
QY 601 CTGCGCCCTCTCTCGCGGCGGAGGCTCACTTGGTCCCGCAGCGCGGCGCTCTCG 660
Db 1850 CTGCGCCCTCTCTCGCGGCGGAGGCTCACTTGGTCCCGCAGCGCGGCGCTCTCG 1909
QY 661 CTGGTCCGGCGCGCGCGCTGCGCGCTGCGCGCGGCTCTTGGAGCAGCGAGG 720
Db 1910 CTGGTCCGGCGCGCGCGCTGCGCGCTGCGCGCGGCTCTTGGAGCAGCGAGG 1969
QY 721 AGCGGGCGCGGCTGCGCTTGGCGGGCGCGGCTCTCAGGATGCGGCTCCCGCGGCTC 780
```

Db 1970 AGCGGGCGCGCTGCGCTTTGCCGGGGCGCGCCCTCCAGGATCGCATCGCGCCGGTC 2029  
QY 781 CGCTGAAGCGCGCGCCCTGCTCGGCGCGAGCGACGACGCGCGACCTCGCCCGG 840  
Db 2030 CGCTGAAGCGCGCGCCCTGCTCGGCGCGAGCGACGACGCGCGACCTCGCCCGG 2089  
QY 841 AGGCTGCAGAGACCGCGGGCGCGCCCTCCCGCTCCCTCTCTCCCGCTCGGCTCTCT 900  
Db 2090 AGGCTGCAGAGACCGCGGGCGCGCCCTCCCGCTCCCTCTCTCCCGCTCGGCTCTCT 2149  
QY 901 CGGCTCTCTGCTCTCAGGGCGCCCTCGCTCCCGCGCGCGAGTCGCTGCGGAGGGC 960  
Db 2150 CGGCTCTCTGCTCTCAGGGCGCCCTCGCTCCCGCGCGCGAGTCGCTGCGGAGGGC 2209  
QY 961 GCCGGCGAGCGCTCTCGGAGGAGCA 986  
Db 2210 GCCGGCGAGCGCTCTCGGAGGAGCA 2235

RESULT 3  
US-09-963-285-5  
; Sequence 5, Application US/09963285  
; Patent No. US20020090707A1  
; GENERAL INFORMATION:  
; APPLICANT: Enerback, Sven  
; APPLICANT: Krook, Katarina  
; APPLICANT: Rondahl, Lena  
; APPLICANT: Wasserman, Wyeth  
; TITLE OF INVENTION: PROMOTER SEQUENCES  
; FILE REFERENCE: 13425-042001  
; CURRENT APPLICATION NUMBER: US/09/963,285  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: SE 0004102-0  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/238,897  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: SE 0003435-5  
; PRIOR FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 6021  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1649)...(438)  
US-09-963-285-5

Query Match 29.4%; Score 290; DB 9; Length 6021;  
Best Local Similarity 64.8%; Pred. No. 2.5e-63;  
Matches 597; Conservative 0; Mismatches 265; Indels 58; Gaps 9;  
QY 23 TTGCTTTTGAATCATTACACCTGGGCGCCCGCCCAATTAATAGGAATCTAATTTTCGGTTCA 82  
Db 1145 TTGCTCTGAACCACTTACACTAGCGCCCGGATTAATTAAGAAATCTAATTTTCGGCTCT.1204  
QY 83 TCACCTCAATTAAT-----AAGAAAATGTCCAGGATCATTTCTACTTACAGGCTCTTT 135  
Db 1205 TCATCCATTAATAATAATAAAAAAATCTCCAGGCTCTTTCTACTTTACAGGCTTTG 1264  
QY 136 GGAGAGATATTTTATCTATTATATCAATTTATTTATTTCAATTTGA-----T 187  
Db 1265 GGGGCAATCTCTGCCCACTTCATCAATTCGATGTATTTTCAACTAACTCTTTT 1324  
QY 188 TTTTTTTACAGAGGAAGTGGCTATCTTTTGTGTTTGGGCGATGTGGGCCCATTCACCAA 247  
Db 1325 TATTTTCCAAAGGAACAGGGTTTTTAATTTTGTCTCTGGACAGTGGTCTCGTAAACAA 1384  
QY 248 AATGTGATCATATAAATAAATTTTAAGATATACT--TTTAAAGTTTTCAGTGA 305  
Db 1385 AATGTGATATAAATAAATTTTAAGATGTAATCAATTTTAAAGTCTCAAGTTA 1444

QY 306 AGACGAGATCGCGCGGAGG-----CCGGGGCGCGGGGTCTTAGAGCCGACGGAT 356  
Db 1445 ACTTGAGCTGGGGGGGGGGAGATCTGGCTAAGAGCATCTGGGTCTTAGAGCCGACGGAT 1504  
QY 357 TCCTCGCTCTCTCGCGCCCGGATTTGGCGCGGACTCCTCTCAGCTGCGGGGTGATTGGCTCA 416  
Db 1505 TCAGGCGCTCTCTCGTTTTGATTGTGCCATCTCTTCGCACTGCGAGATGATTGTGCA 1564  
QY 417 AGTTTCGGGAGGGGGCGTGGCCCGAGAAAAGTAAAACTCGCTTCAGCAAGAAGACTT 476  
Db 1565 AACTTCTCGGAGGGGGCGCGCTTGAAGAAAGTAAAACTCGCTTTGAGCCAGAAGACTT 1624  
QY 477 TTGAAACTTTTCCCAATCCCTAAAGGACTTGGGCTCTTTTCTGGGCTCAGCGGGGCA 536  
Db 1625 TTGAAACTTTTCCCAATCCCTAAAGGACTTGGCTCTTTTCTGGGCTCGGCGCGCA 1684  
QY 537 GCGGCT-CGGACCCCGCGCGCTGACCTCGGGGTGCGGATTCGCTGG-GGGCTTGGAG 594  
Db 1685 GCGCTCGGACCCCTAGCTCGCTGACGCTGGGGGTGCACTTCTCTGGCGGGGCCGAG 1744  
QY 595 AGCCTCTCGCGCCCTCTCTCGCGGGCGGAGGGTCCACCTTGTTCGCCAGCGCGGCG 654  
Db 1745 AGCCTCTCTCTCTTTTCTAGACTCGGAAGG-----CTGGTGTGCTCCA 1791  
QY 655 TCTCCGCTGGGTTCGCGCGCGCGCTGCGCGGCTGCCCGGCTGCCCGCGGGTCTCTGAGCCA 714  
Db 1792 CGGTCGCGCTGGGCTGTGTGCGCGAGCTCAGGCTGCCACCGCCAAAGCGAGAGTGC 1851  
QY 715 GCGAGAGCGGGCGCGGCTGCGCTTGCCTGGGGCGCGCTCCAGGATCCGATCCGC 774  
Db 1852 GCGGCGAGCGGGGCG-----GCCTGCGGTGCACCTTCAGGATCCGATCCGC 1899  
QY 775 CGGTCGCTGAAAGCGCGCGCTCTGCTCGGCGCGGAGCGAGCGACCGCGCACCTCG 834  
Db 1900 CGGTCGCGCTGAACCGGAGCGCGGCTTTCGCG-----CGTGACCGCGAGGCTGCC 1954  
QY 835 CCCGAGGCTGCCAGGAGACGGGGCGCGCTCCCGCTCCCGCTCCCTCTCTCCCGCTCTGG 894  
Db 1955 CCGAGTCGGGGCTGCTGATCGCTCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2014  
QY 895 CTCCTCGCGCTCTCTCGCT 914  
Db 2015 CGCGCGGGCGCGAGTCGGT 2034

RESULT 4  
US-10-116-265-15  
; Sequence 15, Application US/10116265  
; Publication No. US20030077612A1  
; GENERAL INFORMATION:  
; APPLICANT: Erblom, Jonas  
; TITLE OF INVENTION: NEW METHOD  
; FILE REFERENCE: 13425-105001  
; CURRENT APPLICATION NUMBER: US/10/116,265  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: SE 0101218-6  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 60/281,384  
; PRIOR FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 125  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-10-116-265-15

Query Match 12.7%; Score 125; DB 14; Length 125;  
Best Local Similarity 83.2%; Pred. No. 4.6e-22;  
Matches 104; Conservative 21; Mismatches 0; Indels 0; Gaps 0;  
QY 861 CCGCCCTCGCGTCCCTCTCTCCCGCTCTCTCGGCTCTCTCGCTCTCAGG 920



```
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10475
; LENGTH: 51657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(51657)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-10475

Query Match          7.5%; Score 73.6; DB 15; Length 51657;
Best Local Similarity 48.9%; Pred. No. 2.2e-07;
Matches 197; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 512 CTCCTTTTCTGGGCTCAGCGGGGACGCGCTCGAGACCCCGCGCGCTGACCTCGGGGCT 571
Db 17823 CTCCTATTANNNNNNNNGGNNNGGGGGGGCGGGCCNNGCCCGCGGGG 17882

QY 572 GCCGATTGCTGGGGGCTTGAGAGCTCTCTGCGCCCTCTCTGCGCGGGCCGAGGGTCC 631
Db 17883 GCGCGCGGGGCGCGGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 17942

QY 632 ACCTTGCTGCCAGCGCGCGCTCTCGCTGGTCTGCGCGCGCGCGCGCGCGCGCGCT 691
Db 17943 CGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18002

QY 692 GCCGCCCGCGGCTCTGAGAGCGAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCG 751
Db 18003 CCCCGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18060

QY 752 CGCCTCAGAGTCCGATCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
Db 18061 -GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18119

QY 812 GCGACGACGACCGCGCACCTCTGCGCCCGAGGCTGCGAGAGACCGCGGGCGCGCTCC 871
Db 18120 GGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18179

QY 872 GCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 914
Db 18180 GTCCGGCGCCTTTCCCGCTTTCCCGCTTTCCCGCTTTCCCGCTTTCCCGCT 18222

RESULT 8
US-10-154-884B-10475
; Sequence 10475, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
```

```
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10475
; LENGTH: 51657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(51657)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-10475
```

```
Query Match          7.5%; Score 73.6; DB 15; Length 51657;
Best Local Similarity 48.9%; Pred. No. 2.2e-07;
Matches 197; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 512 CTCCTTTTCTGGGCTCAGCGGGGACGCGCTCGAGACCCCGCGCGCTGACCTCGGGGCT 571
Db 17823 CTCCTATTANNNNNNNNGGNNNGGGGGGGCGGGCCNNGCCCGCGGGG 17882

QY 572 GCCGATTGCTGGGGGCTTGAGAGCTCTCTGCGCCCTCTCTGCGCGGGCCGAGGGTCC 631
Db 17883 GCGCGCGGGGCGCGGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 17942

QY 632 ACCTTGCTGCCAGCGCGCGCTCTCGCTGGTCTGCGCGCGCGCGCGCGCGCGCGCT 691
Db 17943 CGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18002

QY 692 GCCGCCCGCGGCTCTGAGAGCGAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCG 751
Db 18003 CCCCGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18060

QY 752 CGCCTCAGAGTCCGATCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
Db 18061 -GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18119

QY 812 GCGACGACGACCGCGCACCTCTGCGCCCGAGGCTGCGAGAGACCGCGGGCGCGCTCC 871
Db 18120 GGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18179

QY 872 GCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 914
Db 18180 GTCCGGCGCCTTTCCCGCTTTCCCGCTTTCCCGCTTTCCCGCTTTCCCGCT 18222
```

```
RESULT 9
US-10-017-161-1403/c
; Sequence 1403, Application US/10017161
; Publication No. US2003014368A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1403
; LENGTH: 1117
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1117)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(917)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (173)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (207)..(366)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (343)..(344)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (351)..(356)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (360)..(361)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (378)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (380)..(383)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (398)..(399)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (401)..(402)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (405)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (431)..(432)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (434)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (437)..(438)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (440)..(444)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (464)..(465)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (468)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (484)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (487)..(497)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (509)..(510)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (525)..(527)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (532)..(533)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (539)..(542)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (548)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (552)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (557)..(561)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (565)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (573)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (577)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (588)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (605)..(610)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (615)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (617)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (619)..(622)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (629)
```

```
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (646)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (648)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (650)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (652)..(653)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (659)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (673)..(675)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (688)..(689)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (694)..(696)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (698)

Query Match 7.48; Score 73; DB 14; Length 1117;
Best Local Similarity 36.08; Pred. No. 3.3e-08;
Matches 157; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

541 CTGGACCCGGCGGCGTGCACCTCGGGGCTGCGGATTCGCTGGGGCTTGGAGAGCCTC 600
858 CNNNNNNNCACNCNCNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 799
601 CTGGGCCCCCTCTCGCGCGGGCGAGGGTCACTTGGTCCCGAGCGCGGCGTCTCCG 660
798 NNCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 739
661 CTGGGTCGCGGCGCGCGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCGAGG 720
738 CCNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 679
721 AGCGGGCGCGCGCTGCGGTTGCGGGGCGCGCCCTCCAGATGCGGATCCCGCGGTC 780
678 CCNNNNCCNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 619
781 CGCTGAAGCGCGGCGCTGCTGGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
618 CNCNCCNNNNNNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 559
841 AGGCTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
558 NNCNCCNCCNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 499
901 CGGCTCTCTGCTCTCAGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
498 CNNNNNNNNNNNNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 439
961 GCCGGCGAGCGGCTCTC 976
438 NNCNCCNCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 423
```

RESULT 10  
US-10-292-798-1141/c  
; Sequence 1141, Application US/10292798  
; Publication No. US20030235833A1  
; GENERAL INFORMATION:  
; APPLICANT: SUWA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA  
; APPLICANT: ABURATANI, HIROYUKI  
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: 084335/166  
; CURRENT APPLICATION NUMBER: US/10/292,798  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 10/017,161  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: JP 2001-246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2070  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1141  
; LENGTH: 1117  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: source  
; FEATURE:  
; LOCATION: (1)..(1117)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(917)  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (170)..(170)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (173)..(173)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (207)..(306)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (343)..(344)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (351)..(356)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (360)..(361)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (378)..(378)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (380)..(383)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (398)..(399)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (401)..(402)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base



```

LOCATION: (405)..(405)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (431)..(432)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (434)..(434)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (437)..(438)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (440)..(444)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (464)..(465)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (468)..(468)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (484)..(484)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (487)..(497)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (509)..(510)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (525)..(527)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (532)..(533)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (539)..(542)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (548)..(548)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (552)..(552)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (557)..(561)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (565)..(565)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (573)..(573)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (577)..(577)

```





FEATURE:  
NAME/KEY: modified base  
LOCATION: (671)..(697)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (704)..(707)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (709)..(736)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (739)..(743)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (745)..(747)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (750)..(750)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (752)..(752)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (754)..(756)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (761)..(761)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (763)..(763)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (818)..(818)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (843)..(843)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (892)..(892)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (897)..(897)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (920)..(920)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (941)..(941)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (993)..(993)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (995)..(995)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:

NAME/KEY: modified base  
LOCATION: (1066)..(1067)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1082)..(1084)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1105)..(1107)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1130)..(1130)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1147)..(1147)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1162)..(1162)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1202)..(1202)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1233)..(1233)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1247)..(1248)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1261)..(1261)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1485)..(1485)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1500)..(1500)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1544)..(1544)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1601)..(1601)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1605)..(1605)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1635)..(1635)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1637)..(1637)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (1654)..(1654)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:53:09 ; Search time 7451.24 seconds  
(without alignments)  
3951.570 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1250\_2235

Perfect score: 986

Sequence: 1 ctgccattccaatccagcgc.....gagccgtctcgaagcagca 986

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hcc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hcc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_eston.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	98.2	10.0	932	29	CNS0072Q
C 2	98.2	10.0	1628	29	CG757066
C 3	96.6	9.8	1201	13	EX340009
C 4	95.4	9.7	1166	12	EX468435

5	94.6	9.6	1406	29	CG756569
6	93.2	9.5	1159	28	CC210735
7	93	9.4	1144	13	EX415926
C 8	92.2	9.4	1046	13	EX415926
C 9	91.8	9.3	1016	13	BQ952554
C 10	91.8	9.3	1023	29	BQ918616
C 11	91.6	9.3	1057	13	AG128304
C 12	91.6	9.3	1273	12	EX349688
C 13	90.4	9.2	716	10	BM562099
14	90.2	9.1	807	13	BB612862
15	90.2	9.1	932	29	BB612862
C 16	90.2	9.1	982	13	BB612862
C 17	89.6	9.1	978	12	EX415111
C 18	89.6	9.1	935	29	CG786339
C 19	89.4	9.1	627	10	CNS006XK
C 20	89.4	9.1	638	13	BB656517
C 21	89.4	9.1	647	10	BB656517
C 22	89.4	9.1	656	13	BB656517
C 23	89.4	9.1	656	13	BB656517
C 24	89.4	9.1	656	13	BB656517
C 25	88.6	9.0	815	14	AK076319
C 26	88.6	9.0	815	14	AK076319
C 27	88	8.9	1341	29	CNS010RY
C 28	87.8	8.9	1341	29	AG030611
C 29	87.8	8.9	1341	29	AG030611
C 30	87.6	8.9	1103	13	CNS006XK
C 31	87.6	8.9	1116	28	CNS015Y4
C 32	87.6	8.9	1201	9	EX403654
C 33	87.4	8.9	356	13	BZ569478
C 34	87.4	8.9	356	13	BZ569478
C 35	87.2	8.8	1598	29	AL563250
C 36	87.2	8.8	982	13	BY310221
C 37	87.2	8.8	1201	13	AG030579
C 38	87.2	8.8	1348	29	EX415111
C 39	87.2	8.8	1798	29	CC190008
C 40	87.2	8.8	1798	29	EX405071
C 41	87	8.8	924	13	CG752544
C 42	87	8.8	1220	29	AG171124
C 43	86.6	8.8	768	12	EX442207
C 44	86.6	8.8	797	29	AG110156
C 45	86.2	8.7	886	28	CG786329
					SEADMC006
					AG030595
					pan trogl
					AZ196424
					SP_1032_A
					BY015578
					BY015578

## ALIGNMENTS

RESULT 1	CNS0072Q/c
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION	BACR14B09 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL086742
VERSION	AL086742.1 GI:4945205
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila.
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
JOURNAL	Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### FEATURES

Location/Qualifiers  
 1. .932  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR14B09"  
 /clone\_lib="RPI-98"  
 /notes="end : 77"

#### ORIGIN

Query Match 10.0%; Score 98.2; DB 29; Length 932;  
 Best Local Similarity 37.4%; Pred. No. 6.5e-07;  
 Matches 171; Conservative 102; Mismatches 181; Indels 3; Gaps 2;

QY 532 GGCACCGCTCGACCCCGCGCTGACCTCGGGCTGCGATTCCTGGGGCTTG 591  
 Db 932 GSGCGCGCGCCGSGGCGSSCGGSCGSCGSCGSCGSCGSCGSCGSCGSCG 873  
 QY 592 GAGAGCTCTCTGCGCCCTCTCTCGCGGCGCGGCTGACCTTGGTCCCGAGCGCG 651  
 Db 872 GCGGCG 813  
 QY 652 GCGTCTCGCTGCG 709  
 Db 812 SSCSSSSCGSCSCGSSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCG 753  
 QY 710 AGCCAGCGAGCGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCTCCAGATGCGGA 769  
 Db 752 SSGCGGCG 693  
 QY 770 TCCGCGCGCGCTGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 829  
 Db 692 GGGCG 633  
 QY 830 CCG 889  
 Db 632 CCG 573  
 QY 890 TCTGGCTCTCTCGCGCTCTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCG 949  
 Db 572 CCGCGVSGCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCG 514  
 QY 950 TGGCGAGGCGCGCGCGCGCGCGCGCTTCGGAAGCAGCA 986  
 Db 513 CCYCGBCCG 477

#### RESULT 2

CG757066/c  
 LOCUS CG757066 1628 bp DNA linear GSS 24-OCT-2003  
 DEFINITION P052-2-A02.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,  
 genomic survey sequence.  
 CG757066  
 VERSION CG757066.1 GI:37985257  
 KEYWORDS GSS.  
 SOURCE Pristionchus pacificus  
 ORGANISM Pristionchus pacificus  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
 Neodiplogasteridae; Pristionchus.  
 REFERENCE 1 (bases 1 to 1628)  
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
 Buntjer,J., van der Weulen,M. and Sommer,R.J.  
 TITLE An integrated physical and genetic map of the nematode Pristionchus  
 pacificus  
 JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)

#### MEDLINE

22835951  
 PUBMED 12884007  
 COMMENT Contact: Sommer RJ  
 Evolutionary Biology  
 Max-Planck-Institute for Developmental Biology  
 Spemannstr. 37-39, Tuebingen D-72076, Germany  
 Tel: 00497071601371  
 Fax: 00497071601498  
 Email: ralf.sommer@tuebingen.mpg.de  
 Class: BAC ends.

#### FEATURES

Location/Qualifiers  
 1. .1628  
 /organism="Pristionchus pacificus"  
 /mol\_type="genomic DNA"  
 /strain="California"  
 /db\_xref="taxon:54126"  
 /clone\_lib="Ppa EcoRI BAC Library"  
 /note="The library was generated by a partial digest of  
 the genomic DNA with EcoRI and cloning into the BAC  
 vector."

#### ORIGIN

Query Match 10.0%; Score 98.2; DB 29; Length 1628;  
 Best Local Similarity 55.3%; Pred. No. 6.5e-07;  
 Matches 231; Conservative 0; Mismatches 178; Indels 9; Gaps 2;  
 QY 531 GGGGCGAGCGCGCTCGGACCCCGCGCGCTGACCTCGGGGCTGCGATTCCTGGGGCTT 590  
 Db 1375 GGGGCG 1316  
 QY 591 GGAGAGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650  
 Db 1315 CGCCCG 1256  
 QY 651 GCGCTCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 710  
 Db 1255 GCG 1198  
 QY 711 GCCAGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 770  
 Db 1197 CG 1138  
 QY 771 CGCGCGCGCGCTGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 830  
 Db 1137 CCGGCG 1085  
 QY 831 CTCGCCCGCGAGGCTGCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 890  
 Db 1084 CG 1025  
 QY 891 CTGGCTCTCTGCGCGCTCTCTGCGCTCTCAGGGCGCGCGCGCGCGCGCG 948  
 Db 1024 CCGGCG 967

#### RESULT 3

CG757066/c  
 LOCUS CG757066 1201 bp mRNA linear EST 02-MAY-2003  
 DEFINITION BX340009 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 clone CS0D1080YF07 3-PRIME, mRNA sequence.  
 CG757066  
 VERSION BX340009  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 Li,W.B., Gruber,C., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
 Buntjer,J., van der Weulen,M. and Sommer,R.J.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage





Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
REFERENCE 1 (bases 1 to 1406)  
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
Buntjer,J., van der Meulen,M. and Sommer,R.J.  
TITLE An integrated physical and genetic map of the nematode Pristionchus  
pacificus  
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)  
MEDLINE 22835951  
PUBMED 12884007  
COMMENT Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: raif.sommer@tuebingen.mpg.de  
Class: BAC ends.

FEATURES  
source  
1..1406  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Ppa EcoRI BAC Library"  
/note="The library was generated by a partial digest of  
the genomic DNA with EcoRI and cloning into the BAC  
vector."

ORIGIN  
Query Match 9.6%; Score 94.6; DB 29; Length 1406;  
Best Local Similarity 49.8%; Pred. No. 2.7e-06;  
Matches 229; Conservative 0; Mismatches 231; Indels 0; Gaps 0;  
QY 523 GGTCTAGCGGCGAGCGCTCGAGCCCGCGCGGTGACCTCGGGGTGCGATTGCT 582  
Db 791 GCGCGGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 850  
QY 583 GGGGGCTTGGAGAGCTCTCTGCGCCCTCTCTCGCGGGCGAGGTTCCACTTGGTCCC 642  
Db 851 CCCCCCG 910  
QY 643 CAGCGCGCGCTCTCCGTGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 702  
Db 911 CCG 970  
QY 703 GTCTTGAGAGCCAGGAGCGGGCGCGCGCGCTTGCGCTTGCGGGCGCGCGCTTCCAGG 762  
Db 971 CG 1030  
QY 763 ATGCCGATCGCGCGGTGCGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822  
Db 1031 GCG 1090  
QY 823 CGCGCACTCTGCGCGCGAGGCTGCGAGAGACCGGGCGCGCGCTTCCCGTCTCCCTCT 882  
Db 1091 CCG 1150  
QY 883 CTCCTCTTGGCTCTCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942  
Db 1151 CG 1210  
QY 943 CAGTCGCTGCGGAGGCGCGCGCGAGCGCTCTCGGAGC 982  
Db 1211 CG 1250

RESULT 6  
CC210735 1159 bp DNA linear GSS 12-MAY-2003  
LOCUS CH261-185P7\_Sp6.2 CH261 Gallus gallus genomic clone CH261-185P7,  
DEFINITION genomic survey sequence.  
CC210735  
ACCESSION

CC210735.1 GI:30529403  
GSS.  
Gallus gallus (chicken)  
SOURCE Gallus gallus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 1159)  
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,  
Warren,W., Graves,T., Mardis,E. and Wilson,R.  
TITLE Gallus gallus BAC End Reads  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Insert Length: 182000 Std Error: 0.00  
Seq primer: Sp6 ATTAGGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 71  
High quality sequence stop: 159.

FEATURES  
source  
1..1159  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-185P7"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/clone\_lib="CH261"  
/note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CH261 Female Chicken library - for library and clone  
ordering information: http://www.chori.org/bacpac"

ORIGIN  
Query Match 9.5%; Score 93.2; DB 28; Length 1159;  
Best Local Similarity 51.3%; Pred. No. 4.6e-06;  
Matches 235; Conservative 0; Mismatches 22; Indels 1; Gaps 1;  
QY 520 CTGGGCTCAGCGGCGAGCGCTCGGACCCCGCGCGCTGACCCCTCGGGGCTCCGATTC 579  
Db 440 CCGCCCG 499  
QY 580 GCTGGGGCTTGAGAGCTCTCTCGCGCGCGCTCTCTCGCGGGCGGAGGCTCCACCTTGGT 639  
Db 500 CCG 559  
QY 640 CCCGAGCGCGCGCTCTCGCTG-GGTCCGCGCGCGCGCGCGCTGCGCGGCTGCGCGCG 698  
Db 560 CCG 619  
QY 699 CCGGCTCTTGAGACCGAGAGCGGGCGCGCTTGCGCTTGCGGGCGCGCGCTC 758  
Db 620 CCNCCCG 679  
QY 759 CAGATCCGATCCGCGCGCTCGCTGAAGCGCGCGCGCGCTGCTCGGCGCGAGCGACGA 818  
Db 680 CCG 739  
QY 819 CGACCGCGCACCTTCGCGCGCGAGGCTGCGAGAGACCGGGCGCGCGCTTCCCGCTCCCC 878  
Db 740 CCG 799  
QY 879 TCCTCTCCCGCTTGAGCTCTCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 938  
Db 800 GCG 859  
QY 939 GCGCGAGTCCGCTGCGGAGCGCGCGCGCGCGCGCGCGCTCTC 976  
Db 860 GCG 897

```
RESULT 7
BX415926/c
LOCUS
DEFINITION BX415926 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YH01
5-PRIME, mRNA sequence.
ACCESSION BX415926
VERSION BX415926.1 GI:30650131
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9016.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008CD01Q1&cluster=9016.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP008CD01Q1.
FEATURES
Location/Qualifiers
source
1..1144
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YH01"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 9.4%; Score 93; DB 13; Length 1144;
Best Local Similarity 39.4%; Pred. No. 5e-06;
Matches 162; Conservative 76; Mismatches 172; Indels 1; Gaps 1;
QY 539 CGCTCGACCCCGCGCGCTGACCTCGGGGCTGCGGATTCGCTGGGGCTTGGAGAGCC 598
Db 1124 CGCGSGSGSGYGGSGCGSGCGSGSGCGGGGCGCGSGCGCGTCCGSGSGGCGC 1065
QY 599 TCCTGCGCCCTCTCTGCGGGGCGAGAGGTCCACCTTG-GTCCCCAGGCGCGGGCTCT 657
Db 1064 SCGCCSSSSCSGSGCSGSGCCSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1005
QY 658 CCCTGAGTCCGCGCGCGCGCGCTCCGCGCTGCGCGCGCGCGCGCGCGCGCGCG 717
Db 1004 SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 945
QY 718 AGAGCGGGCGCGCGCGCTGCGCTTGGCGGGGCGCGCGCTCCAGGATCCGATCCGCG 777
Db 944 CGCGSGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 885
QY 778 GTCCGTGAAGCGCGCGCGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
Db 884 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 825
QY 838 CGAGAGGTGCGAGAGACCGGGGCGCGCGCTCCCGCTCCCGCTCTCTCTCTCTCTCT 897
Db 824 SSGSCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 765
QY 898 TCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 948
```

```
Db 764 SSCSCGGCGCBKTCGCKCKCKGMAAAAMCCCCCCCCCCCCCCCCCCCC 714
RESULT 8
BQ952554/c
LOCUS
DEFINITION BQ952554 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423858
5', mRNA sequence.
ACCESSION BQ952554
VERSION BQ952554.1 GI:22368032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1046)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2603 row: b column: 19
High quality sequence stop: 206.
FEATURES
Location/Qualifiers
source
1..1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6423858"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Query Match 9.4%; Score 92.2; DB 13; Length 1046;
Best Local Similarity 51.8%; Pred. No. 6.9e-06;
Matches 229; Conservative 0; Mismatches 205; Indels 8; Gaps 1;
QY 510 GCCTCTTTTCTGGGCTCAGCGGGGAGCGCTCGAGCCCCCGGGCGCGCTACCTCGGGG 569
Db 859 GCCCGCGNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
QY 570 CTGCGGATTCGCTGGGGGCTTGGAGAGCTCCTCGCGCCCTCTCTCGCGCGCGAGGGT 629
Db 799 CCCCCCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 740
QY 630 CCACCTTTGGTCCCCAGAGCGCGCGCTCTCGCGCTGGGTCCGCGCGCGCGCGCG 689
Db 739 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 680
QY 690 CTGCGCGCGCGGGTCC-----TGAGCCAGCAGAGAGCGGGGCTTGGCGTT 741
Db 679 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 620
QY 742 GCCCGGGCGCGCGCTCCAGGATCCGATCCGCGCGCGCTCCGCTGAAAGCGCGCGCTG 801
Db 619 CCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 560
```

```

QY 802 CTGGGCCCCGAGCAGCAGACCGCGACCTCTGCGCGGAGGCTGCCAGAGACCGGGC 861
Db 559 CCACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500
QY 862 CGGCGCTCTCCGCTCCGCTCTCTCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCT 921
Db 499 CGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 440
QY 922 CCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 943
Db 439 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418

RESULT 9
BQ918616/c 1016 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_8885548 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION IMAGE:6397594 5', mRNA sequence.
ACCESSION BQ918616
VERSION BQ918616.1 GI:22333314
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1895 row: 1 column: 11
High quality sequence stop: 407.
FEATURES
source
1..1016
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="IMAGE:6397594"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 9.3%; Score 91.8; DB 13; Length 1016;
Best Local Similarity 50.2%; Pred. No. 8.1e-06;
Matches 210; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

QY 531 GGGGAGCGCTCGAGACCGCGCGCGCTGACCTCGGGGCTCCCGATTCTGGGGGCTT 590
Db 928 GGGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 869
QY 591 GGAGAGCTCTCTCGCGCGCTCTCTCTCGCGCGCGCGAGGTTCCACCTTGTCTCC 650
Db 868 CGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 809
QY 651 GCGCTCTCGCTGGTTCGCGCGCGCGCGCGCGCTGCCCGCTCCCGCGCGCGGCTCT 710
Db 808 GCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 749
QY 711 GCACGAGAGAGCGGGGCGCGCGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCG 770

```

```

Db 748 CCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 689
QY 771 CCGCCCGGTTCGCTCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 830
Db 688 GGGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 629
QY 831 CTCGCCCGCGAGGCTGCACGAGACCGGGGCGCGCGCGCGCGCGCGCGCGCGCT 890
Db 628 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 569
QY 891 CTGGCTCTCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 948
Db 568 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 511

RESULT 10
AG128304/c 1023 bp DNA linear GSS 04-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-139G20.R, genomic survey sequence.
DEFINITION AG128304
ACCESSION AG128304
VERSION AG128304.1 GI:16657469
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
JOURNAL Totoki, Y., Watanabe, H. and Sakaki, Y.
COMMENT BAC end sequences of library PTB
1
2 (bases 1 to 1023)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..1023
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-139G20.R"
/sex="male"
/call_type="lymphoblast"
/clone_lib="FIB Chimpanzee Male BAC Library"
ORIGIN
Query Match 9.3%; Score 91.8; DB 29; Length 1023;
Best Local Similarity 48.1%; Pred. No. 8.1e-06;
Matches 222; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 524 GCTCAGGGGGGAGCGCGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 583
Db 871 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 812
QY 584 GGGGCTTGGAGAGCGCTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 643
Db 811 GGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 752
QY 644 AGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703

```





```

Db      279  GTCCGCTGAACCCGAGCGCGCGGCTTCTCCGG-----CGTGACCGGAGGCTCCCGC 333
QY      838  CGAGAGCTCCAGGAGACCGGGGCGCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTC 897
Db      334  AGTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 393
QY      898  TC 899
Db      394  TC 395

RESULT 14
BX464554
LOCUS   BX464554
DEFINITION BX464554 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE007YB11
3-PRIME, mRNA sequence.
ACCESSION BX464554
VERSION   BX464554.1 GI:31021646
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 807)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3370.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1DE002ZF03NP1acluster=3370.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1DE002ZF03NP1.

FEATURES             source
    1..807
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DE007YB11"
        /tissue_type="PLACENTA"
        /clone_lib="Homo sapiens PLACENTA"
        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched
double-strand cDNA was digested with NotI and cloned into
the NotI and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match          9.1%; Score 90.2; DB 13; Length 807;
Best Local Similarity 38.8%; Pred. No. 1.5e-05;
Matches 171; Conservative 70; Mismatches 200; Indels 0; Gaps 0;

QY      538  CCCTCGGACCCCGGGCGGTGACCTCCCTCGGGGTGCGGATTCCTCGGGGTGAGAGC 597
Db      263  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 322

QY      598  CTCTGGCCCTCTCTCGCGGGCGAGGGTTCACCTTGGTCCCGAGCGCGGCTCT 557
Db      323  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 382

QY      658  CCCTGGGTCCGGGGCGCGCGCTGCGCGGTGCGCGCGCGGCTCTGAGAGCAGCG 717
Db      383  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 442

QY      718  AGAGAGGGGCGCGGCTGCTGCTTGGCGGGGCGCGCTCTCAGATGCGGATCGCCG 777
Db      443  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 502

```

```

QY      778  GTCCGCTGAAGCGCGGCGCGCTTCTCCGG-----CGTGACCGGAGGCTCCCGC 837
Db      503  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 562
QY      838  CGAGAGGTGCGAGGAGACCGGGGCGCGCTCCCGGTCCCGCTCTCTCCCGCTCTGCTC 897
Db      563  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 622
QY      898  TCTCGCGCTCTCTCGCTCTCAGGGGCGCGCTCTCGTCCCGCGCGGCGAGTCGTCGCGG 957
Db      623  TYBQGYCYCCGSSCCSCCCCGCCCGCCYCCYCCYCCYCCYCCYCCYCCYCCYCCYCCY 978
QY      958  GGCGCGCGGAGCGCTCTCGG 978
Db      683  SKYKSGKXKSCSYBCCCKS 703

RESULT 15
CNS0072Q
LOCUS   CNS0072Q
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066742
VERSION   AL066742
KEYWORDS GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 932)
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
COMMENT   - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oseegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES             source
    1..932
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACR14B09"
        /clone_lib="RPCI-98"
        /note="end : T7"

ORIGIN
Query Match          9.1%; Score 90.2; DB 29; Length 932;
Best Local Similarity 30.7%; Pred. No. 1.5e-05;
Matches 146; Conservative 121; Mismatches 208; Indels 0; Gaps 0;

QY      508  TGGCTCTTTTCTTCTCGGCTCAGCGGGCGCGCTCGGAGCCCGCGCTGACCTCGG 567
Db      452  TGCSTTSBGYCKCKKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 511
QY      568  CGCTGCGGATTCGCTTGGGGGCTTGAGAGCTCTCTCGGCCCTCTCTCGCGGGCGCGG 627

```

```
Db      512 GSGSGSCGTGCGGCGKGCXGCTSKGCTGBSTGCGGGGSKSGSGSGCSBGC GG 571
QY      628 GTCCACCTTGTTCCCGAGCCGCGGCTCTCCGCTGGGTCCGCGGCCGCCCGCTGCGCG 687
Db      572 GGSTKGGSGGGGSSBTKYTBKKGCGGSGSCSGSCRGGGGCGSCGSGGCGCG 631
QY      688 CGCTGCGCGCGCGGGTCTGTGAGCCAGCGAGCGGGCGCGGCTGCGCTTCCCGG 747
Db      632 GSSSGCGSCSCGCGSGGCGCCCGSCSGSGGCSGGSGCGCGCGCCCGCGCC 691
QY      748 GCGCGCGCCCTCCAGGATGCCGATCCGCGCCCGGTCCGCTGAAAGCGCGCGCCCTGCTCGGC 807
Db      692 CSSGSCCCCGCGCGSCCGCGSCSGGCSGCGSCCGSCCGCCCGCCCGSGCS 751
QY      808 CCGAGCGACGACGCGCGCACCTCGCCCCCGAGGCTGCCAGAGAGCGGGGCGCGCCC 867
Db      752 SSSGCGSGGSGSGCGSCGSGCGSGCGSGCGSGCGSGSGSGSGSGSGSGS 811
QY      868 TCCCGTCCCTCTCTCCCGCTCTGGCTCTCTCGGCTCTCTCGCTCTCAGGGCCCCC 927
Db      812 SCGSGSGSCSGSGSGSCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 871
QY      928 TCGTCCCCCGCGCGCGAGTCCGTGCGCGAGGCGCGCGCGCGCGCGCGCGCGCG 982
Db      872 CCSGCGCGCGCGSGSGCGCGCGSGCGSGCGSGCGSGCGSGCGSGSGSGCGCGCG 926
```

Search completed: March 25, 2004, 15:29:54  
Job time : 7456.24 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 07:40:23 ; Search time 12.4532 Seconds  
(without alignments)  
4093.601 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1692\_1703

Perfect score: 12  
Sequence: 1 ggaagtaaaaa 12

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	12	100.0	12	5	ABV43572 Human pro
C 2	12	100.0	19	2	ABQ57116 Human col
C 3	12	100.0	25	8	ACF69945 Photorhab
C 4	12	100.0	28	2	AAV26241 FIV gag g
C 5	12	100.0	28	2	AAV26158 FIV gag-p
C 6	12	100.0	28	3	AAZ34901 FIV gag/p
C 7	12	100.0	28	6	AAZ46912 Feline in
C 8	12	100.0	48	4	AAZ30218 Human SNP
C 9	12	100.0	58	4	AAZ77183 Interfero
C 10	12	100.0	141	3	AAZ98261 Human col
C 11	12	100.0	151	6	ABV199496 Mouse isc
C 12	12	100.0	168	4	ABA75013 Human foe
C 13	12	100.0	168	4	ABA49281 Human liv
C 14	12	100.0	177	5	ABV07368 Human pro
C 15	12	100.0	181	7	ABX60874 Arabidops
C 16	12	100.0	202	5	ABV18575 Human sec
C 17	12	100.0	207	3	AAZ21173 Human bre
C 18	12	100.0	229	7	ABZ71640 Breast sp
C 19	12	100.0	245	4	AAZ25924 Human bre
C 20	12	100.0	254	5	ABA13678 Human ner
C 21	12	100.0	260	5	ABA34006 Human cdn
C 22	12	100.0	269	3	AAZ35275 Arabidops
C 23	12	100.0	273	5	ABV34717 Human pro

ALIGNMENTS

RESULT 1

ABI69613/c  
ID ABI69613 standard; DNA; 12 BP.

XX AC ABI69613;

XX DT 22-FEB-2002 (first entry)

XX DE Oligonucleotide primer SEQ ID NO 369586 for detecting SNP TSC0057728.

XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN WO200177384-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB000713.

XX PR 07-APR-2000; 2000DE-01019173.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;  
WPI; 2001-657177/75.

PT Set of oligonucleotides, useful for diagnosis and cell typing, is designed to detect single-nucleotide polymorphisms and cytosine methylation status.

PS Claim 1; SEQ ID NO 369586; 29pp + Sequence Listing; German.

CC This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The oligonucleotides are used for diagnosis and/or prognosis of cancer and a range of diseases including immune system, gastrointestinal, respiratory, central nervous system, cardiovascular and metabolic disorders. The oligomers are also used for detecting cell type differentiation. ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and ABR00010-ABR20073 represent the oligomers described in the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but

CC was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 12 BP; 1 A; 3 C; 0 G; 8 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 8.7e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GGAAAGTAAAAA 12  
Db 12 GGAAAGTAAAAA 1  
  
RESULT 2  
AAAT74184/c  
ID AAT74184 standard; DNA; 19 BP.  
XX  
AC AAT74184;  
XX  
DT 25-MAR-2003 (revised)  
DT 29-SEP-1997 (first entry)  
XX  
DE Salmonella enteritidis sefB PCR primer.  
XX  
KW Enteropathogenic bacteria; enterobacteria; S. enteritidis; antibody;  
KW polymerase chain reaction; ss.  
XX  
OS Synthetic.  
XX  
PN US5635617-A.  
XX  
PD 03-JUN-1997.  
XX  
PF 26-APR-1994; 94US-00233788.  
XX  
PR 26-APR-1993; 93US-00054452.  
XX  
PA (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.  
XX  
PI Collinson SK, Kay WW, Doran JL;  
XX WPI; 1997-309886/28.  
XX  
PT Isolated Salmonella gene agfa - used for diagnosis of Salmonella or  
PT enteropathogenic bacteria of the Enterobacteria family.  
XX  
PS Example 11; Col 51; 85pp; English.  
XX  
CC The present sequence represents a PCR primer for the amplification of  
CC sefB from Salmonella enteritidis. The nucleic acid produced can be used  
CC to provide diagnostic assays for Salmonella and/or enteropathogenic  
CC bacteria of the family Enterobacteria. It can also be used to provide  
CC proteins and antibodies which can be used for assays. The nucleic acid  
CC sequence can be used to provide probes or primers which can specifically  
CC hybridise to nucleic acid molecules from greater than 99% of Salmonella  
CC strains that are pathogenic to warm-blooded animals relative to nucleic  
CC acid molecules from virtually all other microbial organisms. (Updated on  
CC 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 19 BP; 4 A; 3 C; 2 G; 10 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.7e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GGAAAGTAAAAA 12  
Db 14 GGAAAGTAAAAA 3  
  
RESULT 3  
ACK23905/c  
ACK23905 standard; DNA; 25 BP.  
XX  
AC ACK23905;  
XX  
DT 14-OCT-2003 (first entry)  
XX  
DE Human microarray DNA oligonucleotide SEQ ID NO 123886.  
XX  
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;  
KW genetic variation; biallelic marker; polymorphism; human;  
KW cross-species comparison.  
XX  
OS Homo sapiens.  
XX  
PN US2003104410-A1.  
XX  
PD 05-JUN-2003.  
XX  
PF 15-MAR-2002; 2002US-00098263.  
XX  
PR 16-MAR-2001; 2001US-0276759P.  
XX  
PA (AFFY-) AFFYMETRIX INC.  
XX  
PI Mittmann MP;  
XX  
DR WPI; 2003-567953/53.  
XX  
PT New array of nucleic acid probes, useful for in situ hybridization, in  
PT Southern, Northern or dot-blot hybridization to identify or detect the  
PT sequence or specific mutations of any gene.  
XX  
PS Claim 1; SEQ ID NO 123886; 9pp; English.  
XX  
CC The invention discloses a microarray comprising a plurality of nucleic  
CC acid probes including one of 2,018,500 fully defined sequences, or its  
CC perfect match, perfect mismatch, antisense match or antisense mismatch.  
CC Also disclosed is a method of gene expression analysis. The array is used  
CC in monitoring gene expression levels by hybridisation to a DNA library,  
CC in analysis of genetic variation or in hybridisation of tag-labelled  
CC compounds. The nucleic acid probes are specifically designed for analysis  
CC of at least one target sequence. The method of analysis comprises  
CC hybridising at least one or more nucleic acids to at least two or more  
CC nucleic acid probes and detecting the hybridisation. The nucleic acid  
CC probes are attached to a solid support. The analysis comprises monitoring  
CC gene expression levels, identifying biallelic markers or polymorphisms,  
CC or family members of a gene and a cross-species comparison. Each of the  
CC nucleic acids further comprises a tag sequence. The array of nucleic acid  
CC probes is useful in situ hybridisation, in Southern, Northern or dot-  
CC blot hybridisation to identify or detect the sequence or specific  
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by  
CC primer extensions or in screening cDNA or genomic libraries or subclones  
CC for additional subclones containing segments of DNA that have been  
CC isolated and previously sequenced. The sequence presented is one of the  
CC nucleic acid probes incorporated in the microarray. Note: The sequence  
CC data for this patent can also be obtained in electronic format directly  
CC from USPTO at seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 25 BP; 4 A; 8 C; 4 G; 9 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 8; Length 25;  
Best Local Similarity 100.0%; Pred. No. 8.7e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GGAAAGTAAAAA 12  
Db 24 GGAAAGTAAAAA 13  
  
RESULT 4  
AAV26241/c  
ID AAV26241 standard; DNA; 28 BP.  
XX

```

AC AAV26241;
XX
XX DT 24-JUL-1998 (first entry)
XX
XX DE FIV gag gene PCR primer from WO9804684.
XX
XX KW Swinepox virus; SPV; recombinant; vaccine; immunisation; diagnosis;
XX KW pseudorabies virus; feline immunodeficiency virus; FIV; heartworm;
XX KW Dirofilaria immitis; PCR primer; ss.
XX
XX OS Synthetic.
XX OS Feline immunodeficiency virus.
XX
XX PN WO9804684-A1.
XX
XX PD 05-FEB-1998.
XX
XX PF 25-JUL-1997; 97WO-US012212.
XX
XX PR 25-JUL-1996; 96US-00686968.
XX
XX PA (SYTR ) SYNTRO CORP.
XX
XX CO Cochran MD, Junker DE;
XX
XX PI WPI; 1998-130677/12.
XX
XX DR Recombinant swine pox virus - useful in vaccine for immunising animal
XX PT against swine pox virus.
XX
XX PS Disclosure; Page 76; 473pp; English.
XX
XX CC The present sequence represents a PCR primer from the present invention.
XX CC The present invention specifically describes recombinant swinepox virus
XX CC (SPV) comprising a foreign DNA (I) inserted into a SPV genome which is
XX CC capable of being expressed in a host cell into which the virus is
XX CC introduced, where (I) is inserted into: (a) an EcoRI site within a region
XX CC corresponding to a 3.2 kb subfragment of the HindIII K fragment which
XX CC contains both a HindIII and an EcoRI site, of the SPV genome, and
XX CC optionally (b) an AccI site within a region corresponding to a 3.6 kb
XX CC HindIII to BglII subfragment of the HindIII M fragment. The recombinant
XX CC SPV can be used in a vaccine for immunising an animal against SPV. The
XX CC invention also provides a method for testing a swine to determine whether
XX CC the swine has been vaccinated with the vaccine, particularly containing S
XX CC -SPV-008, or is infected with a naturally occurring wild-type
XX CC pseudorabies virus. Also (I) inserted into recombinant SPV can be used in
XX CC a diagnostic assay, e.g. Feline immunodeficiency virus (FIV) env and gag
XX CC genes and Dirofilaria immitis p39 and 22kd are useful to detect feline
XX CC immunodeficiency caused by FIV and to detect heartworm caused by D.
XX CC immitis respectively
XX
XX SQ Sequence 28 BP; 4 A; 9 C; 4 G; 11 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 12; DB 2; Length 28;
XX Best Local Similarity 100.0%; Pred. No. 8.7e+03;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
DB 27 GGAAAGTAAAAA 16

RESULT 5
AAV26158/c
ID AAV26158 standard; DNA; 28 BP.
XX
XX AC AAV26158;
XX
XX DT 24-JUL-1998 (first entry)
XX
XX DE FIV gag-protease gene PCR primer from WO9804684 Example 44.
XX
XX KW Swinepox virus; SPV; recombinant; vaccine; immunisation; diagnosis;

```

```

KW pseudorabies virus; feline immunodeficiency virus; FIV; heartworm;
XX Dirofilaria immitis; PCR primer; ss.
XX
XX OS Synthetic.
XX OS Feline immunodeficiency virus.
XX
XX PN WO9804684-A1.
XX
XX PD 05-FEB-1998.
XX
XX PF 25-JUL-1997; 97WO-US012212.
XX
XX PR 25-JUL-1996; 96US-00686968.
XX
XX PA (SYTR ) SYNTRO CORP.
XX
XX CO Cochran MD, Junker DE;
XX
XX PI WPI; 1998-130677/12.
XX
XX DR Recombinant swine pox virus - useful in vaccine for immunising animal
XX PT against swine pox virus.
XX
XX PS Example 44; Page 214; 473pp; English.
XX
XX CC The present sequence represents a PCR primer used in an example from the
XX CC present invention. The present invention specifically describes
XX CC recombinant swinepox virus (SPV) comprising a foreign DNA (I) inserted
XX CC into a SPV genome which is capable of being expressed in a host cell into
XX CC which the virus is introduced, where (I) is inserted into: (a) an EcoRI
XX CC site within a region corresponding to a 3.2 kb subfragment of the HindIII
XX CC K fragment which contains both a HindIII and an EcoRI site, of the SPV
XX CC genome, and optionally (b) an AccI site within a region corresponding to
XX CC a 3.6 kb HindIII to BglII subfragment of the HindIII M fragment. The
XX CC recombinant SPV can be used in a vaccine for immunising an animal against
XX CC SPV. The invention also provides a method for testing a swine to
XX CC determine whether the swine has been vaccinated with the vaccine,
XX CC particularly containing S-SPV-008, or is infected with a naturally
XX CC occurring wild-type pseudorabies virus. Also (I) inserted into
XX CC recombinant SPV can be used in a diagnostic assay, e.g. Feline
XX CC immunodeficiency virus (FIV) env and gag genes and Dirofilaria immitis
XX CC p39 and 22kd are useful to detect feline immunodeficiency caused by FIV
XX CC and to detect heartworm caused by D. immitis respectively
XX
XX SQ Sequence 28 BP; 4 A; 9 C; 4 G; 11 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 12; DB 2; Length 28;
XX Best Local Similarity 100.0%; Pred. No. 8.7e+03;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
DB 27 GGAAAGTAAAAA 16

RESULT 6
AAZ34901/c
ID AAZ34901 standard; DNA; 28 BP.
XX
XX AC AAZ34901;
XX
XX DT 28-FEB-2000 (first entry)
XX
XX DE FIV gag/protease gene PCR primer 11/94.10.
XX
XX KW FIV; gag; protease; feline; cat; recombinant virus; vaccine;
XX KW immunomodulator; therapy; PCR; primer; ss.
XX
XX OS Synthetic.
XX OS Feline immunodeficiency virus.
XX
XX PN WO9957295-A1.

```

PD 11-NOV-1999.  
 XX 30-APR-1999; 99WO-US009504.  
 PF 01-MAY-1998; 98US-00071711.  
 XX (SCHE ) SCHERING-PLOUGH LTD.  
 PA (SCHE ) SCHERING-PLOUGH VETERINARY CORP.  
 XX Winslow BJ, Cochran MD;  
 PI WPI; 2000-062155/05.  
 XX Novel recombinant virus useful as immunomodulators, particularly in  
 PT vaccines.  
 XX Disclosure; Page 60; 230pp; English.  
 XX This oligonucleotide represents downstream primer 11/94.10 that was used  
 CC in the PCR amplification of the FIV gag/protease gene. The primer  
 CC introduces a 3' BglII site into the gene. The amplified gene was used in  
 CC the construction of homology vector 904-63.B7, which was used to insert  
 CC foreign DNA into swinepox virus (SPV). The invention relates to a  
 CC recombinant virus, e.g. SPV, that contains at least one foreign nucleic  
 CC acid, inserted into a nonessential genomic region, that encodes feline  
 CC CD28, CD80, CD86 or CTLA-4 protein, or their immunogenic fragments, and  
 CC is expressed when the recombinant virus is introduced into a suitable  
 CC host. The recombinant virus may further comprise a foreign nucleic acid  
 CC encoding an immunogen derived from a feline pathogen such as FIV. It is  
 CC used to enhance an immune response in a feline, particularly as a vaccine  
 XX SQ Sequence 28 BP; 4 A; 9 C; 4 G; 11 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 12; DB 3; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 8.7e+03; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGAAAGTAAAAA 12  
 Db |||||  
 27 GGAAAGTAAAAA 16  
 RESULT 7  
 AAL46912/c  
 ID AAL46912 standard; DNA; 28 BP.  
 XX AAL46912;  
 AC 29-AUG-2003 (revised)  
 DT 07-AUG-2003 (revised)  
 DT 08-AUG-2002 (first entry)  
 DE Feline influenza virus gag/protease gene PCR primer 11/94-10.  
 XX Cat; CD28; CD80; CTLA-4; CD86; immunogen; vaccine; viral infection;  
 KW feline immunodeficiency disease; feline infectious peritonitis;  
 KW feline leukaemia virus; cancer; degenerative disease; autoimmune disease;  
 KW virucide; immunomodulator; cytostatic; immunodeficiency; PCR; primer; ss.  
 XX Viruses.  
 OS US2002051792-A1.  
 XX US2002051792-A1.  
 XX 02-MAY-2002.  
 XX 30-APR-1999; 99US-00303040.  
 XX 01-MAY-1998; 98US-0083870P.  
 PR (WINS/) WINSLOW B J.  
 PA (COCH/) COCHRAN M D.  
 XX Winslow BJ, Cochran MD;  
 PI WPI; 2001-465210/50.  
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
 PT autoimmune diseases and infections.

XX WPI; 2002-415200/44.  
 XX New recombinant virus, useful for immunizing felines to prevent or treat  
 PT feline immunodeficiency virus, comprises foreign nucleic acid encoding  
 PT feline cytotoxic T lymphocyte accessory molecules CD28, CD80, CD86 or  
 PT CTLA-4.  
 XX Disclosure; Page 61; 77pp; English.  
 XX The present invention relates to a recombinant virus comprising at least  
 CC one foreign nucleic acid encoding a protein selected from feline  
 CC cytotoxic T lymphocyte accessory molecules CD28, CD80, CD86 or CTLA-4,  
 CC which is capable of expression when the virus is introduced into an  
 CC appropriate host. The virus can be administered to the feline in order to  
 CC elicit or enhance an immune response to prevent or treat feline  
 CC immunodeficiency disease, feline leukemia, feline infectious peritonitis,  
 CC cancers, degenerative and autoimmune diseases and immunodeficiency. The  
 CC present sequence is a PCR primer described in the exemplification of the  
 CC invention. (Updated on 07-AUG-2003 to correct OS field.) (Updated on 29-  
 CC AUG-2003 to standardise OS field)  
 XX SQ Sequence 28 BP; 4 A; 9 C; 4 G; 11 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 12; DB 6; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 8.7e+03; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGAAAGTAAAAA 12  
 Db |||||  
 27 GGAAAGTAAAAA 16  
 RESULT 8  
 AAL30218/c  
 ID AAL30218 standard; DNA; 48 BP.  
 XX AAL30218;  
 AC 24-JAN-2002 (first entry)  
 DT Human SNP oligonucleotide #3426.  
 DE Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.  
 XX Homo sapiens.  
 OS WO200147944-A2.  
 XX WO200147944-A2.  
 XX 05-JUL-2001.  
 PD 28-DEC-2000; 2000WO-US035498.  
 PF 28-DEC-1999; 99US-0173419P.  
 PR 27-DEC-2000; 2000US-00173419.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 XX WPI; 2001-465210/50.  
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
 PT autoimmune diseases and infections.

PS Claim 1; Page 2369; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic

CC variants of proteins related to amylases, amyloid proteins, angiotensin,

CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,

CC histones, kinases, colony stimulating factors, complement related

CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, g-

CC protein coupled receptors and thioesterases. The present sequence is one

CC such oligonucleotide. The oligonucleotides and the peptides encoded by

CC them may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate expression of the proteins listed above.

CC Disorders that may be prevented, diagnosed and/or treated include

CC multifactorial diseases with a genetic component, such as autoimmune

CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,

CC systemic lupus erythematosus and Grave's disease), inflammation, cancer

CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

CC leukaemia), diseases of the nervous system and an infection of pathogenic

CC organisms

XX

SQ Sequence 48 BP; 7 A; 9 C; 1 G; 31 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 4; Length 48;

Best Local Similarity 100.0%; Pred. NO. 8.6e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTAAAAA 12

DB 30 GGAAGGTAAAAA 19

RESULT 9

AAF77183

ID AAF77183 standard; DNA; 58 BP.

XX

XX

AC AAF77183;

XX

DT 15-JUN-2001 (first entry)

XX

DE Interferon alpha A1 promoter VRE mutant 4PM.

XX

XX Interferon alpha; IFN; A1 promoter; virus responsive element; VRE;

XX interferon regulatory factor; IRF-7; infection; cancer; DNA vaccine; ds.

XX Unidentified.

OS

XX WO200118039-A2.

XX

XX 15-MAR-2001.

PD

XX 01-SEP-2000; 2000WO-GB003367.

FF

XX 03-SEP-1999; 99GB-00020906.

FR

XX (GLAX ) GLAXO GROUP LTD.

XX

XX Pitha-Rowe PM, Yeow WS, Au WC;

PI

XX WPI; 2001-244558/25.

DR

XX

XX Controlling Type I interferon expression level involves modulating level

PT of interferon regulatory factor 7 or its functional analogue which is

PT useful for treating human viral infections or cancer.

XX

XX Example 2; Fig 2; 33pp; English.

PS

XX The present invention relates to controlling the level of Type 1

CC interferon expression by modulating the level of interferon regulatory

CC factor (IRF)-7 or its functional analogue. This is useful for

CC manufacturing a medicament for treating a condition in which an increase

CC in interferon-alpha is beneficial, such as human viral infections or

CC cancer. The method may also be used in the preparation of DNA vaccines

CC for treating viral, bacterial and parasitic infections. The present

CC sequence is the interferon alpha A1 promoter VRE mutant 4PM. This mutant

CC was used to determine precisely where on the VRE IRF-7 acts

XX

SQ Sequence 58 BP; 31 A; 6 C; 14 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 4; Length 58;

Best Local Similarity 100.0%; Pred. NO. 8.6e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTAAAAA 12

DB 12 GGAAGGTAAAAA 23

RESULT 10

AAC98261

ID AAC98261 standard; CDNA; 141 BP.

XX

AC AAC98261;

XX

DT 09-MAR-2001 (first entry)

XX

DE Human colon cancer antigen nucleotide sequence SEQ ID NO:271.

XX

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX identification; cytostatic; cardioactive; neuroprotective; vulnerary;

XX immunomodulatory; muscular; gynaecological; gastrointestinal;

XX nephrotropic; anti-infective; antibacterial; gene therapy; wound;

XX neural disorder; immune system disorder; muscular disorder;

XX reproductive disorder; gastrointestinal disorder; renal disorder;

XX infectious disease; cardiovascular disorder; ss.

XX

OS Homo sapiens.

XX

XX WO200055351-A1.

PN

XX 21-SEP-2000.

PD

XX 08-MAR-2000; 2000WO-US005883.

PF

XX 12-MAR-1999; 99US-0124270P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX

PI Rosen CA, Ruben SM;

PI

XX WPI; 2000-587534/55.

DR

XX P-PSDB; AAB53504.

DR

XX

XX Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

PT disorders such as colon cancer.

XX

XX Claim 1; Page 695; 2104pp; English.

PS

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,

CC called human colon cancer antigens, given in AAB53234 to AAB54006. The

CC human colon cancer antigens can have cytostatic, cardioactive, muscular;

CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and

CC vulnerary, nephrotropic, anti-infective and antibacterial activities, and

CC can be used in gene therapy. The colon cancer antigen polynucleotides,

CC proteins and antibodies to the proteins are useful for the prevention,

CC treatment and diagnosis of colon disorders, such as colon cancer. The

CC polynucleotides may be used in diagnostics and research, such as for

CC chromosome identification, and as hybridisation probes. The proteins may

CC also be used to prevent diseases such as neural disorders, immune system

CC disorders, muscular disorders, reproductive disorders, gastrointestinal

CC disorders, wounds, renal disorders, infectious diseases, and

CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent

CC sequences used in the exemplification of the present invention

XX

SQ Sequence 141 BP; 42 A; 25 C; 28 G; 40 T; 0 U; 6 Other;

Query Match 100.0%; Score 12; DB 3; Length 141;

```
Best Local Similarity 100.0%; Pred. No. 8.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12
DB 79 GGAAGTAAAAA 90

RESULT 11
ABI99496
ID ABI99496 standard; cDNA; 151 BP.
XX AC
XX AC
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:468.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX OS Mus musculus.
XX PN WO200188188-A2.
XX PD
XX PD 22-NOV-2001.
XX XX
XX XX 18-MAY-2001; 2001WO-JP004192.
XX XX 18-MAY-2000; 2000JP-00145977.
XX PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX XX
XX XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX PT expression levels of particular genes defined in the specification or by
XX PT determining the expression profile of a gene group comprising these
XX PT genes.
XX PS Claim 2; Page 1297; 2690pp; English.
XX CC The present invention describes a method for examining ischaemic
XX CC conditions, comprising measuring the expression levels of particular
XX CC genes (I) in a test sample or determining the expression profile of a
XX CC gene group in the sample comprising genes selected from (I). The method
XX CC is useful for examining the ischaemic condition (e.g. compressive
XX CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX CC expression levels of particular genes (ABI99202 to ABI9912, encoding the
XX CC protein sequences in ABB57020 to ABB57374) or by determining the
XX CC expression profile of a gene group comprising these genes. The expression
XX CC levels or expression profiles produced by these genes are used as an
XX CC indicator when screening for ischaemic condition-improving drugs or
XX CC therapeutics for ischaemic diseases. ABI9913 and ABI9914 represent PCR
XX CC primers for a mouse ischaemic condition related sequence, which are used
XX CC in the exemplification of the present invention
XX SQ Sequence 151 BP; 40 A; 36 C; 36 G; 39 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 6; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.6e-03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12
DB 102 GGAAGTAAAAA 113

RESULT 12
ABA75013/c
ID ABA75013 standard; DNA; 168 BP.
XX AC
XX AC
XX DT 25-FEB-2003 (first entry)
XX DE Human liver single exon probe, SEQ ID NO 24271.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
```

```
XX ABA75013;
XX AC
XX DT 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #23318.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD
XX PD 09-AUG-2001.
XX PF
XX PF 30-JAN-2001; 2001WO-US000669.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX XX
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human foetal liver.
XX PS Claim 4; SEQ ID NO 23318; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human foetal liver. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 168 BP; 38 A; 45 C; 37 G; 48 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 8.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12
DB 162 GGAAGTAAAAA 151

RESULT 13
ABS49281/c
ID ABS49281 standard; DNA; 168 BP.
XX AC
XX AC
XX DT 25-FEB-2003 (first entry)
XX DE Human liver single exon probe, SEQ ID NO 24271.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
```

```
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000664.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX Claim 4; SEQ ID NO 24271; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridizes at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS24011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 168 BP; 38 A; 45 C; 37 G; 48 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 8.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12
Db |||||
162 GGAAGTAAAAA 151

RESULT 14
ABV07368.
ID ABV07368 standard; cDNA; 177 BP.
XX
XX ABV07368;
AC
XX 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 7359.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
```

```
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 1191; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX Sequence 177 BP; 41 A; 37 C; 42 G; 48 T; 0 U; 9 Other;

Query Match 100.0%; Score 12; DB 5; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12
Db |||||
120 GGAAGTAAAAA 131

RESULT 15
ABX60874
ID ABX60874 standard; DNA; 181 BP.
XX
XX ABX60874;
AC
XX 26-FEB-2003 (first entry)
XX
XX Arabidopsis thaliana polynucleotide #220.
XX
XX Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
XX genetic modification; environmental stress; disease resistance;
XX fungicide; insecticide; stress tolerance.
XX Arabidopsis thaliana.
XX
XX US2002142319-A1.
XX
XX 03-OCT-2002.
XX
XX 07-AUG-2001; 2001US-00924035.
XX
XX 13-AUG-1999; 99US-0148784P.
XX
XX 11-AUG-2000; 2000US-00638258.
XX
XX (GORL/) GORLACH J.
XX (ANY/) AN Y.
XX (HAMI/) HAMILTON C M.
XX (PRIC/) PRICE J L.
XX (HARG/) HARGISS T R.
XX (YUY/) YU Y.
XX (RAME/) RAMEKA J G.
XX
```

PA (PAGE/) PAGE A.  
PA (MATH/) MATHW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESSNER J F.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JL, Hargiss TR, Yu Y;  
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA;  
XX  
XX WPI; 2003-102509/09.  
XX  
XX Novel Arabidopsis thaliana nucleic acid useful for constructing a  
PT transgenic plant with enhanced disease resistance and enhanced traits of  
PT interest, as probes, and in diagnosis and screening purposes.  
XX  
XX Claim 1; Page 77; 277pp; English.  
XX  
XX The invention relates to Arabidopsis thaliana nucleic acid sequences. The  
CC DNA sequences and the polypeptides they encode are useful for identifying  
CC homologous or related genes, for producing compositions that modulate the  
CC expression or function of the polypeptides, for mapping functional  
CC regions of the protein, in diagnosis, for studying associated  
CC physiological pathways, for genetic manipulation of cells, preferably  
CC plant cells, in screening assays of various plant strains to determine  
CC the strains that are capable of withstanding a particular disease or  
CC environmental stress, for enhancing or inhibiting production of  
CC biosynthetic products in plants and to create genetically modified and  
CC transgenic organisms, such as plant cells and plants. Transgenic plants  
CC are useful for introducing or improving disease resistance and stress  
CC tolerance in plants, screening biologically active agents, such as  
CC fungicides and insecticides, and for identifying factors involved in  
CC biosynthetic pathways of nutritional, commercial or medicinal value.  
CC Sequences ABX6055-ABX61554 represent Arabidopsis thaliana  
CC polynucleotides of the invention  
XX  
SQ Sequence 181 BP; 60 A; 26 C; 25 G; 70 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 7; Length 181;  
Best Local Similarity 100.0%; Pred. No. 8.6e-03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12  
| | | | | | | | | | | | | | | |  
Db 31 GGAAAGTAAAAA 42

Search completed: March 25, 2004, 10:25:09  
Job time : 16.4532 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 10.395 Seconds  
(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1692\_1703

Perfect score: 12  
Sequence: 1 ggaagataaaaa 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/FCI\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/FCIUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	12	100.0	25	14	US-10-098-2638-123886
C 2	12	100.0	28	9	US-09-303-040-80
C 3	12	100.0	141	10	US-09-925-299-271
C 4	12	100.0	141	10	US-09-925-299-271
C 5	12	100.0	149	12	US-10-424-599-472
C 6	12	100.0	150	12	US-10-085-783A-16691
C 7	12	100.0	150	15	US-10-242-535A-16691
C 8	12	100.0	160	12	US-10-424-599-86937
C 9	12	100.0	168	9	US-09-864-761-31340
C 10	12	100.0	172	12	US-10-085-783A-30575
C 11	12	100.0	172	15	US-10-242-535A-30575
C 12	12	100.0	181	9	US-09-924-035A-220
C 13	12	100.0	196	12	US-10-424-599-52857
C 14	12	100.0	206	14	US-10-029-386-17951
C 15	12	100.0	213	12	US-10-424-599-75993

16	12	100.0	221	12	US-10-085-783A-30519	Sequence 30519, A
17	12	100.0	221	15	US-10-242-535A-30519	Sequence 30519, A
C 18	12	100.0	229	14	US-10-082-828A-100	Sequence 100, App
C 19	12	100.0	232	12	US-10-424-599-77849	Sequence 77849, A
20	12	100.0	232	12	US-10-085-783A-8180	Sequence 8180, App
21	12	100.0	232	15	US-10-242-535A-8180	Sequence 8180, App
C 22	12	100.0	235	13	US-10-027-632-178800	Sequence 178800, App
23	12	100.0	250	12	US-10-085-783A-373	Sequence 373, App
24	12	100.0	250	15	US-10-242-535A-373	Sequence 373, App
25	12	100.0	259	12	US-10-424-599-84822	Sequence 84822, A
C 26	12	100.0	264	12	US-10-424-599-95068	Sequence 95068, A
27	12	100.0	273	12	US-10-424-599-7330	Sequence 7330, App
28	12	100.0	276	15	US-10-062-674-77	Sequence 77, App1
29	12	100.0	284	12	US-10-424-599-34915	Sequence 34915, A
30	12	100.0	289	12	US-10-424-599-140768	Sequence 140768, A
C 31	12	100.0	296	12	US-10-424-599-133130	Sequence 133130, A
C 32	12	100.0	302	12	US-10-424-599-23206	Sequence 23206, A
C 33	12	100.0	324	12	US-10-424-599-106240	Sequence 106240, A
C 34	12	100.0	327	12	US-10-424-599-15559	Sequence 15559, A
35	12	100.0	334	12	US-10-424-599-91569	Sequence 91569, A
36	12	100.0	342	12	US-10-424-599-119068	Sequence 119068, A
C 37	12	100.0	355	9	US-09-954-531-829	Sequence 829, App
C 38	12	100.0	355	9	US-09-954-531-1246	Sequence 1246, App
C 39	12	100.0	361	10	US-09-871-161-26	Sequence 26, App1
C 40	12	100.0	361	12	US-10-424-599-105436	Sequence 105436, A
C 41	12	100.0	362	12	US-10-424-599-98384	Sequence 98384, A
C 42	12	100.0	376	12	US-10-424-599-104074	Sequence 104074, A
C 43	12	100.0	384	9	US-09-917-800A-155	Sequence 155, App
C 44	12	100.0	391	9	US-09-960-352-11484	Sequence 11484, A
C 45	12	100.0	396	12	US-10-424-599-119129	Sequence 119129, A

ALIGNMENTS

RESULT 1

US-10-098-263B-123886/c  
; Sequence 123886, Application US/10098263B  
; Publication No. US2003010410A1  
; GENERAL INFORMATION:  
; APPLICANT: Mittman, Michael  
; TITLE OF INVENTION: Human Microarray  
; FILE REFERENCE: 3118.1  
; CURRENT FILING DATE: 2003-01-08  
; PRIOR APPLICATION NUMBER: 60/276,759  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 131066  
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
; SEQ ID NO 123886  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-098-263B-123886

Query Match 100.0%; Score 12; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.1e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12

Db 24 GGAAGTAAAAA 13

RESULT 2

US-09-303-040-80/c  
; Sequence 80, Application US/09303040  
; Patent No. US20020051792A1  
; GENERAL INFORMATION:  
; APPLICANT: Winslow, Barbara J.  
; APPLICANT: Cochran, Mark D.  
; TITLE OF INVENTION: Recombinant Virus Expressing Foreign DNA Encoding  
; TITLE OF INVENTION: Feline CD80, Feline CD96, Feline CD28, Feline CTLA-4 or

RESULT 4  
US-09-925-299-271  
; Sequence 271, Application US/09925299  
; Publication No. US20030040617A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 271  
; LENGTH: 141  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (12)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (13)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (14)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (25)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (39)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (141)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-271

Query Match 100.0%; Score 12; DB 9; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.1e+04; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
Db 27 GGAAGTAAAAA 16

RESULT 3  
US-09-925-299-271  
; Sequence 271, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 271  
; LENGTH: 141  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (12)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (13)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (14)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (25)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (39)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (141)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-271

Query Match 100.0%; Score 12; DB 9; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
Db 79 GGAAGTAAAAA 90

Query Match 100.0%; Score 12; DB 10; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
|||||  
Db 79 GGAAGTAAAAA 90

RESULT 5  
US-10-424-599-472/c  
; Sequence 472, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 472  
; LENGTH: 149  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(149)  
; OTHER INFORMATION: unsure at all n locations

FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_100426C.1  
US-10-424-599-472

Query Match 100.0%; Score 12; DB 12; Length 149;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 138 GGAAGTAAAAA 127

RESULT 6  
US-10-085-783A-16691  
; Sequence 16691, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 16691  
; LENGTH: 150  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-16691

Query Match 100.0%; Score 12; DB 12; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 130 GGAAGTAAAAA 141

RESULT 7  
US-10-242-535A-16691  
; Sequence 16691, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 16691  
; LENGTH: 150  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-16691

Query Match 100.0%; Score 12; DB 15; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 130 GGAAGTAAAAA 141

RESULT 8  
US-10-424-599-86937/c  
; Sequence 86937, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 86937  
; LENGTH: 160  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(160)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_49514C.1  
US-10-424-599-86937

Query Match 100.0%; Score 12; DB 12; Length 160;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 32 GGAAGTAAAAA 21

RESULT 9  
US-09-864-761-31340/c  
; Sequence 31340, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aesmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667

;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,697  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 31340  
;; LENGTH: 168  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL35797.3  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4  
;; OTHER INFORMATION: NT HIT: D49387.1, EVALUATE 5.00e-90  
;; OTHER INFORMATION: SWISSPROT HIT: Q14914, EVALUATE 1.00e-26  
;; OTHER INFORMATION: EST\_HUMAN HIT: BE884686.1, EVALUATE 7.00e-90  
US-09-864-761-31340

Query Match 100.0%; Score 12; DB 9; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
Db 162 GGAAGTAAAAA 151

RESULT 10  
US-10-085-783A-30575  
; Sequence 30575, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 30575  
; LENGTH: 172  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-30575

Query Match 100.0%; Score 12; DB 12; Length 172;

Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
Db 69 GGAAGTAAAAA 80

RESULT 11  
US-10-242-535A-30575  
; Sequence 30575, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 30575  
; LENGTH: 172  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-30575

Query Match 100.0%; Score 12; DB 15; Length 172;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
Db 69 GGAAGTAAAAA 80

RESULT 12  
US-09-924-035A-220  
; Sequence 220, Application US/09924035A  
; Patent No. US20020142319A1  
; GENERAL INFORMATION:  
; APPLICANT: Grilach, Jtn  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2011US  
; CURRENT APPLICATION NUMBER: US/09/924,035A  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,784  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 220  
; LENGTH: 181  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-924-035A-220

Query Match 100.0%; Score 12; DB 9; Length 181;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
Db 31 GGAAGTAAAAA 42

RESULT 13  
US-10-424-599-52857/c  
; Sequence 52857, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 52857  
; LENGTH: 196  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_18742C.1  
US-10-424-599-52857

Query Match 100.0%; Score 12; DB 12; Length 196;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 23 GGAAGTAAAAA 12

RESULT 14  
US-10-029-386-17951/c  
; Sequence 17951, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 17951  
; LENGTH: 206  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR22 36.0  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 14  
; OTHER INFORMATION: NT HIT: gi15718688, EVALUE 2.00e-78  
; OTHER INFORMATION: SWISSPROT HIT: P97351, EVALUE 2.00e-27  
; OTHER INFORMATION: EST\_HUMAN HIT: BF908569.1, EVALUE 1.00e-80  
US-10-029-386-17951

Query Match 100.0%; Score 12; DB 14; Length 206;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 49 GGAAGTAAAAA 38

RESULT 15  
US-10-424-599-75993  
; Sequence 75993, Application US/10424599

Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 75993  
; LENGTH: 213  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_39636C.1  
US-10-424-599-75993

Query Match 100.0%; Score 12; DB 12; Length 213;  
Best Local Similarity 100.0%; Pred. No. 1.2e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12  
| | | | | | | | | |  
Db 90 GGAAGTAAAAA 101

Search completed: March 25, 2004, 15:52:47  
Job time: 11.395 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:53:09 ; Search time 90.6845 Seconds  
(without alignments)  
3951.570 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_1692\_1703

Perfect score: 12

Sequence: 1 ggaagtaaaaa 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	37	28	AZ949109
2	12	100.0	65	13	BQ078148
3	12	100.0	66	29	CG427224
4	12	100.0	67	28	BH850672

5	12	100.0	73	29	CG400241
6	12	100.0	78	14	CD919593
7	12	100.0	81	28	CC049447
8	12	100.0	85	29	EX547641
9	12	100.0	94	29	CG846984
10	12	100.0	102	9	AV832401
11	12	100.0	105	29	CG063748
12	12	100.0	108	29	CG466155
13	12	100.0	111	14	CF084097
14	12	100.0	112	9	A1509743
15	12	100.0	113	29	AL942319
16	12	100.0	114	28	BH078181
17	12	100.0	117	12	BI495080
18	12	100.0	119	29	CG466358
19	12	100.0	122	28	BH86770
20	12	100.0	126	13	BU865302
21	12	100.0	126	13	EX297444
22	12	100.0	127	9	AA729081
23	12	100.0	129	10	AW924842
24	12	100.0	131	10	BE955882
25	12	100.0	131	28	CC048069
26	12	100.0	134	28	AZ043161
27	12	100.0	136	9	AU256208
28	12	100.0	136	13	BU833252
29	12	100.0	137	29	AL759355
30	12	100.0	138	12	BI452728
31	12	100.0	142	28	AZ889185
32	12	100.0	145	29	CG570474
33	12	100.0	146	9	AV354395
34	12	100.0	148	12	BM750390
35	12	100.0	149	10	BE290481
36	12	100.0	149	12	BI188234
37	12	100.0	149	12	BI191170
38	12	100.0	150	9	AV309354
39	12	100.0	150	13	BX766472
40	12	100.0	151	9	AA103768
41	12	100.0	152	10	BF823246
42	12	100.0	152	28	AZ67766
43	12	100.0	156	28	BZ487651
44	12	100.0	157	9	AA827341
45	12	100.0	157	10	BB073490

ALIGNMENTS

RESULT 1  
AZ949109  
LOCUS 37 bp DNA linear GSS 27-APR-2001  
DEFINITION 2M0212K19F Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
clone UUGC2M0212K19 F, genomic survey sequence.  
ACCESSION AZ949109  
VERSION AZ949109.1 GI:13820336  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 37)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112 USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0212 row: K column: 19  
 Seq primer: CATTGTAAACGACGCGCAGT  
 Class: plasmid ends  
 High quality sequence stop: 37.  
 Location/Qualifiers

## FEATURES

source

1. 37  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC2M0212K19"  
 /sex="female"  
 /lab\_host="E. coli strain XL10-Gold, Ti-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC2M library"  
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 Kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GII4732114[gb]/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 100.0%; Score 12; DB 28; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 10 GGAAGTAAAAA 21

## RESULT 2

BQ078148  
 LOCUS f85501.y1 Zebrafish C32 14 somite embryo Danio rerio cDNA clone  
 DEFINITION IMAGE:5816785 5', mRNA sequence.

ACCESSION BQ078148  
 VERSION BQ078148.1 GI:19907185  
 KEYWORDS EST.  
 SOURCE Danio rerio (zebrafish)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

## REFERENCE

1 (bases 1 to 65)  
 AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcevic, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

## TITLE

WashU Zebrafish EST Project 1998

## JOURNAL

Unpublished (1998)

## COMMENT

Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu  
 cDNA Library construction by: Joe Barnes and Steve Johnson. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: RessourcenZentrumPrimardatenbank, Berlin, Germany (web address: www.rzpd.de)  
 Seq primer: T3 ET from Amersham.  
 Location/Qualifiers

## FEATURES

source

1. 65  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /strain="C32"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:5816785"  
 /tissue\_type="embryo, 14 somite"  
 /lab\_host="DH10B"  
 /clone\_lib="Zebrafish C32 14 somite embryo"  
 /notes="Vector: pAMP1; Site 1: EcoRI; Site 2: NotI; First strand cDNA synthesis was primed using oligo-dT on magnetic beads with an additional primer 5'-ggcgccgctaatagcactacta-taggg-3'. Second strand synthesis was a 3-cycle PCR using the primers 5'-ggcgccgctaatagcactactag-3', 5'-aagcagtggttaacacgcagactctt-tttttttttt-3', cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggcgccgctaatagcactactag-3', 5'-aagcagtggttaacacgcag-3'. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers 5'-caucacuaacgagcgctaatagcactactaggg-3' and 5'-cuacacuaacgagcgctaatagcactactag-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."

## ORIGIN

Query Match 100.0%; Score 12; DB 13; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTAAAAA 12  
 |||||  
 Db 21 GGAAGTAAAAA 32

## RESULT 3

CG427224  
 LOCUS 01S0726-07A1-H10 UniformMu Mutail Library Zea mays genomic clone  
 DEFINITION 01S0726-07A1-H10, genomic survey sequence.

ACCESSION CG427224  
 VERSION CG427224.1 GI:34735704

## KEYWORDS

GSS.

## SOURCE

Zea mays

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

1 (bases 1 to 66)

## AUTHORS

Latshaw, S., Tan, B.-C., Settles, A.M. and McCarty, D.R.

## TITLE

Sequence tagged transposon insertions from the UniformMu maize

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Donald R. McCarty  
 Plant Molecular and Cellular Biology Program  
 University of Florida  
 PO 110690 Gainesville, FL 32611-0690, USA  
 Tel: 352-392-1928 x322  
 Email: drmc@ufl.edu  
 Sequence flanking probable Mu insertion site in UniformMu line:  
 01S0726-07, Primer set: A  
 Class: transposon insertion site.  
 Location/Qualifiers

```

source
1. .66
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone="01S0726-07A1-H10"
/clone_lib="UniformMu MutAIL Library"
/notes="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN
Query Match 100.0%; Score 12; DB 29; Length 66;
Best Local Similarity 100.0%; Pred. NO. 3.3e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
Db 35 GGAAAGTAAAAA 46

RESULT 4
LOCUS BH850672 67 bp DNA linear GSS 13-JUN-2002
DEFINITION SALK_071670.53.75.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_071670.53.75.x, genomic
survey sequence.
ACCESSION BH850672
VERSION BH850672.1 GI:21421543
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 67)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1. .67
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_071670.53.75.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN

Query Match 100.0%; Score 12; DB 29; Length 66;
Best Local Similarity 100.0%; Pred. NO. 3.3e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
Db 35 GGAAAGTAAAAA 46

RESULT 4
LOCUS BH850672 67 bp DNA linear GSS 13-JUN-2002
DEFINITION SALK_071670.53.75.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_071670.53.75.x, genomic
survey sequence.
ACCESSION BH850672
VERSION BH850672.1 GI:21421543
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 67)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1. .67
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_071670.53.75.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN

Query Match 100.0%; Score 12; DB 29; Length 73;
Best Local Similarity 100.0%; Pred. NO. 3.2e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
Db 35 GGAAAGTAAAAA 46

RESULT 5
LOCUS CG400241 73 bp DNA linear GSS 02-SEP-2003
DEFINITION O1S0561-04A1-H12 UniformMu MutAIL Library Zea mays genomic clone
O1S0561-04A1-H12, genomic survey sequence.
ACCESSION CG400241
VERSION CG400241.1 GI:34399125
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 73)
AUTHORS Latehaw,S., Tan,B.-C., Settles,A.M. and McCarty,D.R.
TITLE Sequence tagged transposon insertions from the UniformMu maize
population
JOURNAL Unpublished (2003)
COMMENT Contact: Donald R. McCarty
Plant Molecular and Cellular Biology Program
University of Florida
PO 110690 Gainesville, FL 32611-0690, USA
Tel: 352-392-1928 x322
Email: drmc@ufl.edu
Sequence flanking probable Mu insertion site in UniformMu line:
O1S0561-04, Primer set: A
Class: transposon insertion site.
Location/Qualifiers
1. .73
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone="O1S0561-04A1-H12"
/clone_lib="UniformMu MutAIL Library"
/notes="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN

Query Match 100.0%; Score 12; DB 29; Length 73;
Best Local Similarity 100.0%; Pred. NO. 3.2e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12
Db 35 GGAAAGTAAAAA 46

RESULT 6
LOCUS CD919593 78 bp mRNA linear EST 15-JUL-2003
DEFINITION G608.113005F010910 G608 Triticum aestivum cDNA clone G608113005,
mRNA sequence.
ACCESSION CD919593
VERSION CD919593.1 GI:32767357
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum

```



Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 78)  
Genoplante, a major partnership french program in plant genomics  
Unpublished (2003)  
Contact: Genoplante  
Genoplante  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplante' (<http://www.genoplante.com>  
and <http://genoplante-info.infobio.gen.fr>).

FEATURES

source  
1..78  
Location/Qualifiers  
/organism="Triticum aestivum"  
/mol\_type="rRNA"  
/cultivar="recital"  
/db\_xref="taxon:4565"  
/clone="G608113005"  
/tissue\_type="grain (608 degrees per day after  
pollination)"  
/clone\_lib="G608"

ORIGIN

Query Match 100.0%; Score 12; DB 14; Length 78;  
Best Local Similarity 100.0%; Pred. No. 3.1e+05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12  
Db 60 GGAAAGTAAAAA 71

RESULT 7

CC049447  
LOCUS  
DEFINITION  
CC049447  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

81 bp DNA linear GSS 20-OCT-2003  
O1S-458-3-7c012-G02, genomic survey sequence.

CC049447.1 GI:29464338

GSS.

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 81)

Latschaw.S., Tan,B.-C., Settles,A.M. and McCarty,D.R.

Sequence tagged transposon insertions from the UniformMu maize population

Unpublished (2003)

Contact: Donald R. McCarty

Plant Molecular and Cellular Biology Program

University of Florida

PO 110690 Gainesville, FL 32611-0690, USA

Tel: 352-392-1928 x322

Email: [drmc@ufl.edu](mailto:drmc@ufl.edu)

Sequence flanking probable Mu insertion site in UniformMu line:

O1S-458-3

Class: transposon insertion site.

FEATURES  
source

1..81  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="W22 (ACR, bz1-m9)"  
/cultivar="UniformMu"  
/db\_xref="taxon:4577"  
/clone="O1S-458-3-7c012-G02"  
/clone\_lib="UniformMu MutAIL Library"  
/notes="Vector: TOPO-PCR4; DNA flanking Mu transposon

insertions in Mu inactive lines were extracted from the UniformMu maize population by the thermo asymmetric interlaced PCR (TAIL) protocol using primers specific for the Mu terminal inverted repeat and a set of 16 arbitrary primers. Amplicons were size enriched using Sepharose 400 spin columns and cloned into the TOPO PCR4 vector."

ORIGIN

Query Match 100.0%; Score 12; DB 28; Length 81;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTAAAAA 12  
Db 35 GGAAAGTAAAAA 46

RESULT 8

EX547641/C

LOCUS

DEFINITION  
EX547641  
Arabidopsis thaliana T-DNA flanking sequence GK-546F08-020630,  
genomic survey sequence.

ACCESSION

EX547641

VERSION

EX547641.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H. and Weisshaar,B.

TITLE

A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

Unpublished

JOURNAL

REFERENCE

AUTHORS

Strizhov,N., Li,Y., Rosso,M., Dekker,K. and Weisshaar,B.

TITLE

A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

Unpublished

JOURNAL

REFERENCE

AUTHORS

Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.

TITLE

Direct Submission

Submitted (01-JUL-2003)

Weisshaar B., Max-Planck-Institut fuer

Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50823, Germany

This sequence is recovered from the left border of the T-DNA. It

indicates an insertion close to or within gene At3G02040. The

sequences are generated at the MPI for Plant Breeding Research in

the context of the GABI-Kat project. GABI-Kat is part of the German

Plant Genomics program designated 'GABI'. Information on line

availability can be found at:

<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

Location/Qualifiers

1..85

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/strain="Columbia 0"

/db\_xref="taxon:3702"

/clone="GK-546F08-020630"

/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"

/notes="PCR was performed on DNA from Arabidopsis thaliana

plants (T1) which were transformed with the T-DNA from

vector pAC161. The lines contain one or more T-DNA

insertions. The DNA fragment(s) resulting from the PCR

were directly sequenced to determine the genomic sequence

flanking the insertion. Sequences displaying significant

similarity to the A. thaliana nuclear genome sequence were

processed for submission. T-DNA derived sequences were

removed"

ORIGIN

```

Query Match      100.0%; Score 12; DB 29; Length 85;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
    |||||
Db 68 GGAAAGTAAAAA 57

RESULT 9
CG846984
LOCUS      CG846984      94 bp      DNA      linear      GSS 17-NOV-2003
DEFINITION CG846984 01S0554-03C1-C07 UniformMu MutAIL Library Zea mays genomic clone
O1S0554-03C1-C07, genomic survey sequence.
ACCESSION  CG846984
VERSION     CG846984.1  GI:38351227
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 94)
AUTHORS     Letshaw,S., Tan,B.-C., Settles,A.M. and McCarty,D.R.
TITLE       Sequence tagged transposon insertions from the UniformMu maize
            population
JOURNAL     Unpublished (2003)
COMMENT     Contact: Donald R. McCarty
            Plant Molecular and Cellular Biology Program
            University of Florida
            PO 110690 Gainesville, FL 32611-0690, USA
            Tel: 352-392-1928 x322
            Email: drmc@ufl.edu
            Sequence flanking probable Mu insertion site in UniformMu line:
            O1S0554-03, Primer set: C
            Class: transposon insertion site.
FEATURES    Location/Qualifiers
             1..94
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="W22 (ACR, bz1-m9)"
                /cultivar="UniformMu"
                /db_xref="taxon:4577"
                /clone="O1S0554-03C1-C07"
                /clone_lib="UniformMu MutAIL Library"
                /notes="Vector: TOPO-PCR4; DNA flanking Mu transposon
            insertions in Mu inactive lines were extracted from the
            UniformMu maize population by the thermo asymmetric
            interlaced PCR (TAIL) protocol using primers specific for
            the Mu terminal inverted repeat and a set of 16 arbitrary
            primers. Amplicons were size enriched using Sepharose 400
            spin columns and cloned into the TOPO PCR4 vector."

ORIGIN
Query Match      100.0%; Score 12; DB 29; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
    |||||
Db 35 GGAAAGTAAAAA 46

RESULT 10
AV832401
LOCUS      AV832401      102 bp      mRNA      linear      EST 22-JUN-2001
DEFINITION AV832401 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
vulgare leaves vegetative stage Hordeum vulgare subsp. vulgare cDNA
clone baak12a18, mRNA sequence.
ACCESSION  AV832401
VERSION     AV832401.1  GI:14524490
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 105)
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B/3"

```

```

ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            1 (bases 1 to 102)
            Sato,K.
            Barley EST sequencing project in NIG and Okayama Univ
            Unpublished (2001)
            Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kzsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
            submission:
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES    Location/Qualifiers
             1..102
                /organism="Hordeum vulgare subsp. vulgare"
                /mol_type="mRNA"
                /cultivar="Akashinriki"
                /sub_species="vulgare"
                /db_xref="taxon:112509"
                /clone="baak12a18"
                /tissue_type="leaves"
                /dev_stage="vegetative stage"
                /clone_lib="K. Sato unpublished cDNA library: Hordeum
            vulgare subsp. vulgare leaves vegetative stage"

ORIGIN
Query Match      100.0%; Score 12; DB 9; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.7e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12
    |||||
Db 28 GGAAAGTAAAAA 39

RESULT 11
CG063748
LOCUS      CG063748      105 bp      DNA      linear      GSS 19-AUG-2003
DEFINITION CG063748 ZM 0.6.1.0 KB Zea mays genomic clone ZMMBTa0537F14,
genomic survey sequence.
ACCESSION  CG063748
VERSION     CG063748.1  GI:33935928
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 105)
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B/3"
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other_GSSs: PUTAH31TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: IR
            Class: sheared ends.
            Location/Qualifiers
             1..105
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B/3"

```

```

/db_xref="taxon:4577"
/clone="ZMBETA0537F14"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 100.0%; Score 12; DB 29; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.7e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
    |||
Db 46 GGAAGTAAAAA 57

RESULT 12
CG466155
LOCUS
DEFINITION 01S0579-09A1-H10 UniformMu MuTAIL Library Zea mays genomic clone
ACCESSION 01S0579-09A1-H10, genomic survey sequence.
VERSION CG466155
KEYWORDS GSS.
SOURCE CG466155.1 GI:36335395
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 108)
AUTHORS Lathaw,S., Tan,B.-C., Settles,A.M. and McCarty,D.R.
TITLE Sequence tagged transposon insertions from the UniformMu maize
JOURNAL population
COMMENT Contact: Donald R. McCarty
Plant Molecular and Cellular Biology Program
University of Florida
PO 110690 Gainesville, FL 32611-0690, USA
Tel: 352-392-1928 x322
Email: drmcufl.edu
Sequence flanking probable Mu insertion site in UniformMu line:
01S0579-09, Primer set: A
Class: transposon insertion site.

FEATURES
source
1..108
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone="01S0579-09A1-H10"
/clone_lib="UniformMu MuTAIL Library"
/note="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN
Query Match 100.0%; Score 12; DB 29; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.7e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
    |||
Db 74 GGAAGTAAAAA 85

RESULT 13
CF084097
LOCUS

```

```

DEFINITION QHL21J23.YG.ab1 QH_L sunflower H.paradoxus (salt stress) Helianthus
Paradoxus cDNA clone QHL21J23, mRNA sequence.
ACCESSION CF084097
VERSION CF084097.1 GI:33123140
KEYWORDS EST.
SOURCE Helianthus paradoxus
ORGANISM Helianthus paradoxus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 111)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and sunflower ESTs from the Compositae Genome Project
JOURNAL http://compgenomics.ucdavis.edu/
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig5265, see http://cgdb.ucdavis.edu/
for details.
Plate: QHL21 row: J column: 23.

FEATURES
source
1..111
Location/Qualifiers
/organism="Helianthus paradoxus"
/mol_type="mRNA"
/db_xref="taxon:73304"
/clone="QHL21J23"
/lab_host="E.coli"
/clone_lib="QH_L sunflower H.paradoxus (salt stress)"
/note="Vector: pGEM-T; The library was constructed from
four different sources (seedling, root, leaf and flower)
of RNA from a single genotype. cDNAs were pooled and
cloned into a high-copy vector pGEM-T. Details of library
construction can be obtained at http://cgdb.ucdavis.edu/"

ORIGIN
Query Match 100.0%; Score 12; DB 14; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTAAAAA 12
    |||
Db 43 GGAAGTAAAAA 54

RESULT 14
AI509743 112 bp mRNA linear EST 12-MAR-1999
LOCUS IMAGE1365287.5, mRNA sequence.
DEFINITION vx23d04.v1 Soares thymus 2NDMT Mus musculus cDNA clone
ACCESSION AI509743
VERSION AI509743.1 GI:4408648
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 112)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the

```

IMAGE Consortium (info@image.llnl.gov) for further information.  
 MG1:667839  
 This read is a RESSEQUENCE of a previously sequenced mouse clone  
 This read has been verified (found to hit its original self in the  
 correct orientation)

Seq primer: -40RP from Gibco.

#### FEATURES

Location/Qualifiers  
 1. .112  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1265287"  
 /sex="male"  
 /tissue\_type="Thymus"  
 /dev\_stage="4 weeks"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares thymus 2N6MT"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTTTTTTTT  
 3']; double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. RNA  
 provided by Dr. Bertrand Jordan. Library went through two  
 rounds of normalization, and was constructed by Bento  
 Soares and M.Fatima Bonaldo."

#### ORIGIN

Query Match 100.0%; Score 12; DB 9; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12  
 |||||  
 Db 65 GGAAAGTAAAAA 76

#### RESULT 15

AL942319  
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-265A11-014998, 113 bp DNA linear GSS 24-OCT-2002  
 DEFINITION genomic survey sequence.

ACCESSION AL942319

VERSION AL942319.1 GI:24398917

KEYWORDS GSS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H.  
 and Weisshaar,B.

2 A pipeline for automated high-throughput generation of FSTs  
 (flanking sequence tags) from Arabidopsis thaliana T-DNA  
 transformed lines

Unpublished

JOURNAL

REFERENCE

2 Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.  
 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)

for flanking sequence tag based reverse Genetics

Unpublished

3 (bases 1 to 113)

Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.

Direct Submission

Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer  
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence is recovered from the left border of the T-DNA. It  
 indicates an insertion within the locus defined by clone MSD24. The  
 sequences are generated at the MPI for Plant Breeding Research in  
 the context of the GABI-Kat project. GABI-Kat is part of the German

Plant Genomics program designated 'GABI'. Information on line  
 availability can be found at:  
 http://www.mpiz-koeln.mpg.de/GABI-Kat/.

#### FEATURES

source

Location/Qualifiers  
 1. .113  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="GK-265A11-014998"  
 /clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
 /note="PCR was performed on DNA from Arabidopsis thaliana  
 plants (T1) which were transformed with the T-DNA from  
 vector PAC161. The lines contain one or more T-DNA  
 insertions. The DNA fragment(s) resulting from the PCR  
 were directly sequenced to determine the genomic sequence  
 flanking the insertion. Sequences displaying significant  
 similarity to the A. thaliana nuclear genome sequence were  
 processed for submission. T-DNA derived sequences were  
 removed"

#### ORIGIN

Query Match 100.0%; Score 12; DB 29; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAAGTAAAAA 12

|||||

Db 1 GGAAAGTAAAAA 12

Search completed: March 25, 2004, 15:29:59

Job time : 95.6845 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:53:09 ; Search time 68.0134 Seconds

(without alignments)  
3951.570 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_223\_231

Perfect score: 9

Sequence: 1 acaaatgtt 9

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1493109076 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_ges\_hum:\*

18: em\_ges\_inv:\*

19: em\_ges\_pln:\*

20: em\_ges\_vrt:\*

21: em\_ges\_fun:\*

22: em\_ges\_mam:\*

23: em\_ges\_mus:\*

24: em\_ges\_pro:\*

25: em\_ges\_rod:\*

26: em\_ges\_pbg:\*

27: em\_ges\_vrl:\*

28: gb\_ges1:\*

29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	100.0	22	28	AZ465994
C 2	9	100.0	24	28	AZ481958
C 3	9	100.0	24	28	AZ563647
4	9	100.0	25	29	TA212F07P

C 5	9	100.0	26	28	BH864357
C 6	9	100.0	26	29	CC888324
C 7	9	100.0	31	29	AL764318
C 8	9	100.0	33	9	AU255936
C 9	9	100.0	34	9	AA930856
C 10	9	100.0	34	9	AU255210
C 11	9	100.0	34	28	BZ597035
C 12	9	100.0	37	28	AZ827214
C 13	9	100.0	40	28	AZ401018
C 14	9	100.0	41	28	AZ463986
C 15	9	100.0	42	29	AL946600
C 16	9	100.0	43	9	AI931104
C 17	9	100.0	43	28	AZ617367
C 18	9	100.0	44	28	AZ818897
C 19	9	100.0	44	28	BH906176
C 20	9	100.0	46	9	AU257598
C 21	9	100.0	46	28	BH864657
C 22	9	100.0	47	9	AU254247
C 23	9	100.0	47	28	BH792299
C 24	9	100.0	48	29	CC795092
C 25	9	100.0	50	9	AU102686
C 26	9	100.0	50	9	AU103894
C 27	9	100.0	50	28	BH911549
C 28	9	100.0	51	12	BM859286
C 29	9	100.0	51	28	AZ307595
C 30	9	100.0	52	28	AZ826355
C 31	9	100.0	52	28	BH861183
C 32	9	100.0	54	28	AZ777486
C 33	9	100.0	55	28	AZ303648
C 34	9	100.0	55	28	AZ625335
C 35	9	100.0	56	9	AA600694
C 36	9	100.0	56	14	T99557
C 37	9	100.0	57	14	CA333590
C 38	9	100.0	57	28	AZ417527
C 39	9	100.0	57	28	BZ290162
C 40	9	100.0	58	9	AA600618
C 41	9	100.0	58	14	D45774
C 42	9	100.0	58	28	BH790740
C 43	9	100.0	58	29	BX651171
C 44	9	100.0	60	29	AL751735
C 45	9	100.0	61	9	AI629352

#### ALIGNMENTS

#### RESULT 1

AZ465994/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ465994 22 bp DNA linear GSS 04-OCT-2000  
IM0276015F Mouse 10kb plasmid UNGC1M library Mus musculus genomic  
clone UNGC1M0276015 F, genomic survey sequence.

AZ465994 1 GI:10624119

GSS.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 22)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhauser, A. and Wright, D. Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0276 row: O column: 15  
 Seq primer: CCGTGTAAACACGCGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 22.  
 Location/Qualifiers

# FEATURES

source  
 1..22  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0276015"  
 /sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
 |||||  
 Db 19 ACAATGTT 11

RESULT 2  
 AZ481958/c 24 bp DNA linear GSS 04-OCT-2000  
 LOCUS  
 DEFINITION IM0306G07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0306G07 R, genomic survey sequence.

ACCESSION AZ481958  
 VERSION AZ481958.1 GI:10643023  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 24)  
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0306 row: G column: 07  
 Seq primer: CACACAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 24.  
 Location/Qualifiers

# FEATURES

source  
 1..24  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0306G07"  
 /sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
 |||||  
 Db 9 ACAATGTT 1

RESULT 3  
 AZ663647 24 bp DNA linear GSS 14-DEC-2000  
 LOCUS  
 DEFINITION IM0543G20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0543G20 F, genomic survey sequence.

ACCESSION AZ663647  
 VERSION AZ663647.1 GI:11800793  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 24)  
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0543 row: G column: 20  
 Seq primer: CGTTGTAACACGACGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 24.

# FEATURES

Location/Qualifiers  
 1..24  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0543G20"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (G14732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
 |||||  
 Db 11 ACAAATGTT 19

RESULT 4  
 TA212F07P 25 bp DNA linear GSS 13-DEC-2000  
 LOCUS  
 DEFINITION  
 T. brucei sheared genomic DNA clone 212f07, forward sequence,  
 genomic survey sequence.

ACCESSION AL479789  
 VERSION AL479789.1 GI:11845470  
 KEYWORDS  
 SOURCE GSS.

ORGANISM  
 Trypanosoma brucei  
 Trypanosoma brucei  
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma

REFERENCE 1 (bases 1 to 25)  
 Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,  
 Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,  
 Melville,S.E., Rajandream,M.A. and Barrell,B.G.  
 Direct Submission

TITLE Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing  
 JOURNAL project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,  
 Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and  
 ml@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR),  
 Rockville, MD. Genomic DNA isolated from a cloned population of  
 Trypanosoma brucei (TREU927/4 GUTAT 10.1) was mechanically sheared  
 to give a tight size distribution (4 kb). The v + i method used for the library construction is  
 described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In  
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Barrell, Oxford University Press, 1999).  
 Email: nelsayed@tigr.org  
 Details of T. brucei sequencing at the Sanger Centre are available  
 at http://www.sanger.ac.uk/Projects/T\_brucei/.

# FEATURES

Location/Qualifiers  
 1..25  
 /organism="Trypanosoma brucei"  
 /mol\_type="genomic DNA"  
 /strain="TREU927"  
 /db\_xref="taxon:5691"  
 /clone="212f07"

# ORIGIN

Query Match 100.0%; Score 9; DB 29; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
 |||||  
 Db 17 ACAAATGTT 25

RESULT 5  
 BH864357/c

LOCUS  
 DEFINITION  
 SALK\_095812 Arabidopsis thaliana TDNA insertion lines Arabidopsis  
 thaliana genomic clone SALK\_095812, genomic survey sequence.

ACCESSION BH864357  
 VERSION BH864357.1 GI:22100255  
 KEYWORDS  
 SOURCE GSS.

ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana

REFERENCE 1 (bases 1 to 26)  
 Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,  
 Gadriinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,  
 Shinn,P., Zimmerman,J. and Ecker,J.R.  
 A Sequence-Indexed Library of Insertion Mutations in the  
 Arabidopsis Genome  
 Unpublished (2001)

AUTHORS  
 CONTACT: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu

JOURNAL

COMMENT  
 This is single pass sequence recovered from the left border of  
 TDNA. This sequence lies within an annotated intron of At3g53580.  
 Class: TDNA tagged.

FEATURES  
 Location/Qualifiers  
 1..26  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="SALK\_095812"  
 /clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
 /note="PCR was performed on Arabidopsis thaliana lines  
 each of which contains one or more TDNA insertion  
 elements. The resultant fragment for each line was  
 directly sequenced to determine the genomic sequence at  
 the site of insertion. Details of the protocols used can  
 be found at http://signal.salk.edu/tdna\_protocols.html"

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 ACAAATGTT 9
Db      19 ACAAATGTT 11

RESULT 6
CC888324
LOCUS   26 bp      DNA      linear      GSS 31-JUL-2003
DEFINITION
Arabidopsis thaliana T-DNA insertion lines
survey sequence.
ACCESSION
CC888324.1 GI:33364841
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 26)
REFERENCE
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
T-DNA.
Class: T-DNA tagged.
FEATURES
source
1..26
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_151694.27.60.x"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more T-DNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN
Query Match 100.0%; Score 9; DB 29; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACAAATGTT 9
Db      2 ACAAATGTT 10

RESULT 7
AL764318/c
LOCUS   31 bp      DNA      linear      GSS 18-JUN-2002
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence GK-124C06-012813,
genomic survey sequence.
ACCESSION
AL764318
VERSION
AL764318.1 GI:21517141
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 26)
REFERENCE
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
T-DNA.
Class: T-DNA tagged.
FEATURES
source
1..26
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_151694.27.60.x"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more T-DNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN
Query Match 100.0%; Score 9; DB 29; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACAAATGTT 9
Db      21 ACAAATGTT 13

RESULT 8
AU255936/c
LOCUS   33 bp      mRNA      linear      EST 25-APR-2002
DEFINITION
AU255936 3'-directed mouse cDNA library Mus musculus cDNA clone
BE0006850 3', mRNA sequence.
ACCESSION
AU255936
VERSION
AU255936.1 GI:20319148
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 33)
REFERENCE
Kato,K. and Matoba,R.
Generation of expressed sequence tags from mouse brain
Unpublished (2002)
CONTACT: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 31)
REFERENCE
Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
for flanking sequence tag based reverse Genetics
Unpublished
3 (bases 1 to 31)
REFERENCE
Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
Direct Submission
Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It
indicates an insertion close to or within gene Atlg70560. The
sequences are generated at the MPI for Plant Breeding Research in
the context of the GABI-Kat project. GABI-Kat is part of the German
Plant Genomics program designated 'GABI'. Information on line
availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
Location/Qualifiers
1..31
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-124C06-012813"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/notes="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector PAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

ORIGIN
Query Match 100.0%; Score 9; DB 29; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACAAATGTT 9
Db      21 ACAAATGTT 13

RESULT 8
AU255936/c
LOCUS   33 bp      mRNA      linear      EST 25-APR-2002
DEFINITION
AU255936 3'-directed mouse cDNA library Mus musculus cDNA clone
BE0006850 3', mRNA sequence.
ACCESSION
AU255936
VERSION
AU255936.1 GI:20319148
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 33)
REFERENCE
Kato,K. and Matoba,R.
Generation of expressed sequence tags from mouse brain
Unpublished (2002)
CONTACT: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology

```



8916-5 Takayama, Ikoma, Nara 630-0101, Japan  
Tel: 81-743-72-5581  
Fax: 81-743-72-5589  
Email: kkatob@b.aist-nara.ac.jp,  
URL: http://love2.aist-nara.ac.jp/BED/index.html.  
Location/Qualifiers

## FEATURES

1. .33  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="BED0006850"  
/tissue\_type="brain"  
/clone\_lib="3'-directed mouse cDNA library"

## source

## ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGTT 9  
|||||  
Db 19 ACAATGTT 11

## RESULT 9

## AA930856/c

## LOCUS

DEFINITION v271c06.s1 Soares mammary gland NbMWG Mus musculus cDNA clone  
IMAGE:1331914 3' similar to SW:UB5C HUMAN P47986  
UBIQUITIN-CONJUGATING ENZYME E2-17 KD 3', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

AA930856 34 bp mRNA linear EST 23-APR-1998  
IMAGE:1331914 3' similar to SW:UB5C HUMAN P47986  
UBIQUITIN-CONJUGATING ENZYME E2-17 KD 3', mRNA sequence.  
AA930856 1 GI:3080244  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 34)  
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:691458

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28m13 rev2 BT from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers

## FEATURES

## source

1. .34  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1331914"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH103"  
/clone\_lib="Soares mammary gland NbMWG"  
/notes="Organ: mammary gland; Vector: pRT3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;

Site 2: Eco RI: 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5,  
TGTTACCAATCTGAAGTGGAGCGCGCAATGGTTTTTTTTTTTTTTTTTT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pRT3 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M. Fatima  
Bonaldo."

## ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGTT 9  
|||||  
Db 22 ACAATGTT 14

## RESULT 10

## AU255210

## LOCUS

DEFINITION AU255210 3'-directed mouse cDNA library Mus musculus cDNA clone  
BED0004806 3', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

AU255210 34 bp mRNA linear EST 25-APR-2002  
BED0004806 3', mRNA sequence.  
AU255210 1 GI:20317731  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 34)  
Kato, K. and Matoba, R.  
Generation of expressed sequence tags from mouse brain  
Unpublished (2002)  
Contact: Kikuya Kato  
Graduate School of Biological Sciences  
Nara Institute of Science and Technology  
8916-5 Takayama, Ikoma, Nara 630-0101, Japan  
Tel: 81-743-72-5581  
Fax: 81-743-72-5589  
Email: kkatob@b.aist-nara.ac.jp,  
URL: http://love2.aist-nara.ac.jp/BED/index.html.  
Location/Qualifiers

## FEATURES

## source

## 1. .34

## /organism="Mus musculus"

## /mol\_type="mRNA"

## /db\_xref="taxon:10090"

## /clone="BED0004806"

## /tissue\_type="brain"

## /clone\_lib="3'-directed mouse cDNA library"

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## QY

## Db

## RESULT 11

## BZ597035

## LOCUS

## DEFINITION

BZ597035 34 bp DNA linear GSS 07-JAN-2003  
SALK 099290.19.00.n Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_099290.19.00.n, genomic  
survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 34)  
AUTHORS Alonso,J.M., Leishe,T.J., Barajas,P., Chen,H., Cheuk,R.,  
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,  
Shinn,P., Zimmerman,J. and Ecker,J.R.  
TITLE A Sequence-Indexed Library of Insertion Mutations in the  
Arabidopsis Genome  
JOURNAL Unpublished (2001)  
COMMENT Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 5379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of  
TDNA.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .34  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="SALK 099290.19.00.n"  
/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
/note="PCR was performed on Arabidopsis thaliana lines  
each of which contains one or more TDNA insertion  
elements. The resultant fragment for each line was  
directly sequenced to determine the genomic sequence at  
the site of insertion. Details of the protocols used can  
be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

ORIGIN  
Query Match 100.0%; Score 9; DB 28; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05; Length 34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAATGTT 9  
| | | | |  
Db 13 ACAATGTT 21

RESULT 12  
AZ827214/c  
LOCUS 2M0103H02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC2M0103H02 R, genomic survey sequence.  
ACCESSION AZ827214  
VERSION AZ827214.1 GI:12997122  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 37)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 34)  
AUTHORS Alonso,J.M., Leishe,T.J., Barajas,P., Chen,H., Cheuk,R.,  
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,  
Shinn,P., Zimmerman,J. and Ecker,J.R.  
TITLE A Sequence-Indexed Library of Insertion Mutations in the  
Arabidopsis Genome  
JOURNAL Unpublished (2001)  
COMMENT Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 5379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of  
TDNA.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .37  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0103H02"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 Kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWB42 (GI4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 9; DB 28; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e+05; Length 37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAATGTT 9  
| | | | |  
Db 35 ACAATGTT 27

RESULT 13  
AZ401018  
LOCUS 1M0167K04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0167K04 R, genomic survey sequence.  
ACCESSION AZ401018  
VERSION AZ401018.1 GI:10516092  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plats: 0103 row: H column: 02  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 37.  
Location/Qualifiers  
1. .37  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0103H02"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 Kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWB42 (GI4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
Query Match 100.0%; Score 9; DB 28; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e+05; Length 37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAATGTT 9  
| | | | |  
Db 35 ACAATGTT 27

RESULT 13  
AZ401018  
LOCUS 1M0167K04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0167K04 R, genomic survey sequence.  
ACCESSION AZ401018  
VERSION AZ401018.1 GI:10516092  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0167 row: K column: 04  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 40.

# FEATURES

Location/Qualifiers  
 1. .40  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0167K04"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (G14732114|GB|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
 Db 4 ACAATGTT 12

# RESULT 14

AZ463986/c  
 LOCUS AZ463986 41 bp DNA linear GSS 04-OCT-2000  
 DEFINITION 1M0273J07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 Clone UUGC1M0273J07 F, genomic survey sequence.

ACCESSION AZ463986  
 VERSION 1  
 KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 41)

REFERENCE 1 (bases 1 to 41)  
 Authors Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
 Niederhausen, A. and Wright, D., Weiss, R.,  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0273 row: J column: 07  
 Seq primer: CGTTGTAACAGCGGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 41.

# FEATURES

Location/Qualifiers  
 1. .41  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0273J07"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (G14732114|GB|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# ORIGIN

Query Match 100.0%; Score 9; DB 28; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGTT 9  
 Db 37 ACAATGTT 29

# RESULT 15

AL946600  
 LOCUS AL946600 42 bp DNA linear GSS 24-OCT-2002  
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-297G01-015513,  
 genomic survey sequence.

ACCESSION AL946600  
 VERSION 1  
 KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1

REFERENCE 1  
 Authors Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H.  
 and Weisshaar, B.  
 A pipeline for automated high-throughput generation of ESTs  
 (flanking sequence tags) from Arabidopsis thaliana T-DNA  
 transformed lines

Unpublished  
 2

# JOURNAL

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

TITLE Direct Submission  
 JOURNAL Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer  
 Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
 COMMENT This sequence is recovered from the left border of the T-DNA. It  
 indicates an insertion close to or within gene Atg15370. The  
 sequences are generated at the MPI for Plant Breeding Research in  
 the context of the GABI-Kat project. GABI-Kat is part of the German  
 Plant Genomics program designated 'GABI'. Information on line  
 availability can be found at:  
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES  
 source  
 1..42  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="GK-297G01-015513"  
 /note="PCR was performed on DNA from Arabidopsis thaliana  
 plants (T1) which were transformed with the T-DNA from  
 vector pAC161. The lines contain one or more T-DNA from  
 insertions. The DNA fragment(s) resulting from the PCR  
 were directly sequenced to determine the genomic sequence  
 flanking the insertion. Sequences displaying significant  
 similarity to the A. thaliana nuclear genome sequence were  
 processed for submission. T-DNA derived sequences were  
 removed"

ORIGIN

Query Match 100.0%; Score 9; DB 29; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGTT 9  
 |||||  
 Db 13 ACAATGTT 21

Search completed: March 25, 2004, 15:30:04  
 Job time : 73.0134 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 7.79622 Seconds  
(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_223\_231

Perfect score: 9  
Sequence: 1 acaaatgtt 9

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_5/ptodata/1/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	9	100.0	20	US-10-370-860A-14
2	9	100.0	22	Sequence 14, Appl
3	9	100.0	22	Sequence 84, Appl
4	9	100.0	22	Sequence 57, Appl
5	9	100.0	22	Sequence 296, App
6	9	100.0	23	Sequence 4249, App
7	9	100.0	23	Sequence 363, App
8	9	100.0	24	Sequence 8, Appl
9	9	100.0	25	Sequence 3149, App
10	9	100.0	25	Sequence 8452, App
11	9	100.0	25	Sequence 15021, A
12	9	100.0	25	Sequence 15022, A
13	9	100.0	25	Sequence 16053, A
14	9	100.0	25	Sequence 16054, A
15	9	100.0	25	Sequence 22881, A
			25	Sequence 22882, A

16	9	100.0	25	14	US-10-098-263B-23315	Sequence 23315, A
17	9	100.0	25	14	US-10-098-263B-23316	Sequence 23316, A
18	9	100.0	25	14	US-10-098-263B-23317	Sequence 23317, A
19	9	100.0	25	14	US-10-098-263B-23318	Sequence 23318, A
20	9	100.0	25	14	US-10-098-263B-40465	Sequence 40465, A
21	9	100.0	25	14	US-10-098-263B-44365	Sequence 44365, A
22	9	100.0	25	14	US-10-098-263B-47225	Sequence 47225, A
23	9	100.0	25	14	US-10-098-263B-49716	Sequence 49716, A
24	9	100.0	25	14	US-10-098-263B-53713	Sequence 53713, A
25	9	100.0	25	14	US-10-098-263B-70095	Sequence 70095, A
26	9	100.0	25	14	US-10-098-263B-73711	Sequence 73711, A
27	9	100.0	25	14	US-10-098-263B-74037	Sequence 74037, A
28	9	100.0	25	14	US-10-098-263B-74038	Sequence 74038, A
29	9	100.0	25	14	US-10-098-263B-81119	Sequence 81119, A
30	9	100.0	25	14	US-10-098-263B-81120	Sequence 81120, A
31	9	100.0	25	14	US-10-098-263B-84101	Sequence 84101, A
32	9	100.0	25	14	US-10-098-263B-84393	Sequence 84393, A
33	9	100.0	25	14	US-10-098-263B-89011	Sequence 89011, A
34	9	100.0	25	14	US-10-098-263B-92451	Sequence 92451, A
35	9	100.0	25	14	US-10-098-263B-92418	Sequence 92418, A
36	9	100.0	25	14	US-10-098-263B-95977	Sequence 95977, A
37	9	100.0	25	14	US-10-098-263B-105011	Sequence 105011, A
38	9	100.0	25	14	US-10-098-263B-105012	Sequence 105012, A
39	9	100.0	25	14	US-10-098-263B-105173	Sequence 105173, A
40	9	100.0	25	14	US-10-098-263B-110291	Sequence 110291, A
41	9	100.0	25	14	US-10-098-263B-125932	Sequence 125932, A
42	9	100.0	25	14	US-10-098-263B-126835	Sequence 126835, A
43	9	100.0	25	14	US-10-098-263B-126836	Sequence 126836, A
44	9	100.0	26	10	US-09-754-853A-1012	Sequence 1012, Ap
45	9	100.0	26	14	US-10-362-552-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-10-370-860A-14  
; Sequence 14, Application US/10370860A  
; Publication No. US20030219791A1  
; GENERAL INFORMATION:  
; APPLICANT: NOVO NORDISK A/S  
; APPLICANT: Wahl, Philip  
; APPLICANT: No. US20030219791A1by, Peder Lisby  
; APPLICANT: Grondahl, Christian  
; APPLICANT: Stennicke, Vibeke Westphal  
; TITLE OF INVENTION: A Transducer of Mas Signalling  
; FILE REFERENCE: 6475.200-US  
; CURRENT APPLICATION NUMBER: US/10370,860A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/231,670  
; PRIOR FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: PA 2002 00277  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 09/934,948  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: PCT/DK01/00550  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: PA 2000 01259  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-370-860A-14

Query Match 100.0%; Score 9; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
 Db ||||| 16  
 8 ACAAATGTT 16

## RESULT 2

US-09-989-643-84/c  
 ; Sequence 84, Application US/09989643  
 ; Publication No. US20030049636A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bergeron, Michel G.  
 ; APPLICANT: Picard, Francois J.  
 ; APPLICANT: Ouellette, Marc  
 ; APPLICANT: Roy, Paul H.  
 ; TITLE OF INVENTION: Species-Specific, Genus-Specific and Universal DNA  
 ; TITLE OF INVENTION: Probes and Amplification Primers to Rapidly Detect and  
 ; TITLE OF INVENTION: Identify Common Bacterial and Fungal Pathogens and  
 ; TITLE OF INVENTION: Associated Antibiotic Resistance Genes from  
 ; FILE REFERENCE: 12287.29  
 ; CURRENT APPLICATION NUMBER: US/09/989,643  
 ; CURRENT FILING DATE: 2001-11-20  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/297,539  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-09  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/743,637  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1996-11-04  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 84  
 ; LENGTH: 22  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA  
 US-09-989-643-84

Query Match 100.0%; Score 9; DB 10; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+04; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 ACAAATGTT 9  
 Db ||||| 13  
 13 ACAAATGTT 5

## RESULT 3

US-09-864-029-57/c  
 ; Sequence 57, Application US/09864029  
 ; Publication No. US20030082174A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Majumder, Kumud  
 ; APPLICANT: Tchernev, Velizar T.  
 ; APPLICANT: Groses, William M.  
 ; APPLICANT: Szekeres Jr., Edward S.  
 ; APPLICANT: Alsbrook II, John P.  
 ; APPLICANT: Burgess, Catherine E.  
 ; APPLICANT: Shinkets, Richard A.  
 ; APPLICANT: Taupier Jr., Raymond J.  
 ; APPLICANT: Casman, Stacie J.  
 ; APPLICANT: Gangolli, Esha A.  
 ; APPLICANT: MacDougall, John R.  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Smithson, Glennda  
 ; TITLE OF INVENTION: NOVEL GPCR-LIKE PROTEINS AND NUCLEIC ACIDS ENCODING  
 ; TITLE OF INVENTION: SAME  
 ; FILE REFERENCE: 21402-022  
 ; CURRENT APPLICATION NUMBER: US/09/864,029  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 60/206,757  
 ; PRIOR FILING DATE: 2000-05-24  
 ; PRIOR APPLICATION NUMBER: 60/214,372  
 ; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 60/219,786  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 60/207,020  
 ; PRIOR FILING DATE: 2000-05-25  
 ; PRIOR APPLICATION NUMBER: 60/220,593  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: 60/239,542  
 ; PRIOR FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: 60/256,402  
 ; PRIOR FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: 60/271,645  
 ; PRIOR FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/274,809  
 ; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 60/275,590  
 ; PRIOR FILING DATE: 2001-03-13  
 ; NUMBER OF SEQ ID NOS: 59  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 57  
 ; LENGTH: 22  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:  
 ; OTHER INFORMATION: oligonucleotide primer  
 US-09-864-029-57

Query Match 100.0%; Score 9; DB 10; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+04; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 ACAAATGTT 9  
 Db ||||| 21  
 21 ACAAATGTT 13

## RESULT 4

US-10-114-270-296  
 ; Sequence 296, Application US/10114270  
 ; Publication No. US20040030110A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Guo, Xiaojia  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Liu, Ziaohong  
 ; APPLICANT: Gusev, Vladimir Y.  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Pena, Carol E.A.  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Burgess, Catherine E.  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Shinkets, Richard A.  
 ; APPLICANT: Gangolli, Esha A.  
 ; APPLICANT: Taupier Jr., Raymond J.  
 ; APPLICANT: Casman, Stacie J.  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Liete, Mario W.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: MacDougall, John R.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-322C

```

; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 296
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Reverse Primer
US-10-114-270-296

```

```

Query Match 100.0%; Score 9; DB 12; Length 22;
Best Local Similarity 100.0%; Pred.No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAATGTT 9

Db 5 ACAATGTT 13

## RESULT 5

```

US-10-032-585-4249/c
; Sequence 4249, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boosey
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4249
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-4249

```

```

Query Match 100.0%; Score 9; DB 14; Length 23;
Best Local Similarity 100.0%; Pred.No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAATGTT 9

Db 10 ACAATGTT 2

## RESULT 6

US-10-210-130-363/c

```

; Sequence 363, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vermet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Garlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: Dipippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369

```

```
/ SOFTWARE: CuraSeqList version 0.1
/ SEQ ID NO 363
/ LENGTH: 23
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Primer/Probe
US-10-210-130-363

Query Match          100.0%; Score 9; DB 15; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9
Db 23 ACAAATGTT 15

RESULT 7
US-10-452-591-8/c
/ Sequence 8, Application US/10452591
/ Publication No. US20040014144A1
/ GENERAL INFORMATION:
/ APPLICANT: MADSEN, Mogens Winkel
/ APPLICANT: OLSEN, Lone Stengelsboj
/ APPLICANT: FJORDING, Marianne Scheel
/ TITLE OF INVENTION: Method of Screening for Substances Acting on MSK1
/ FILE REFERENCE: 3893-0218P
/ CURRENT APPLICATION NUMBER: US/10/452,591
/ CURRENT FILING DATE: 2003-06-03
/ PRIOR APPLICATION NUMBER: PCT/DK00/00505
/ PRIOR FILING DATE: 2000-09-13
/ PRIOR APPLICATION NUMBER: US 60/159,092
/ PRIOR FILING DATE: 1999-10-13
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 8
/ LENGTH: 24
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic primer for N-terminal part of MSK1
US-10-452-591-8

Query Match          100.0%; Score 9; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9
Db 20 ACAAATGTT 12

RESULT 8
US-09-969-373-3149
/ Sequence 3149, Application US/09969373
/ Patent No. US20020133852A1
/ GENERAL INFORMATION:
/ APPLICANT: Eifertz, Roger J.
/ APPLICANT: Hauge, Brian M.
/ TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
/ FILE REFERENCE: 38-10(52679)A
/ CURRENT APPLICATION NUMBER: US/09/969,373
/ CURRENT FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US 09/754,853
/ PRIOR FILING DATE: 2001-01-05
/ PRIOR APPLICATION NUMBER: US 09/760,427
/ PRIOR FILING DATE: 2001-01-13
/ PRIOR APPLICATION NUMBER: US 09/855,768
/ PRIOR FILING DATE: 2001-05-15
/ NUMBER OF SEQ ID NOS: 4593
/ SEQ ID NO 3149
/ LENGTH: 25

/ TYPE: DNA
/ ORGANISM: Glycine max
US-09-969-373-3149

Query Match          100.0%; Score 9; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9
Db 14 ACAAATGTT 22

RESULT 9
US-10-098-263B-8462/c
/ Sequence 8462, Application US/10098263B
/ Publication No. US20030104410A1
/ GENERAL INFORMATION:
/ APPLICANT: Mittman, Michael
/ TITLE OF INVENTION: Human Microarray
/ FILE REFERENCE: 3118.1
/ CURRENT APPLICATION NUMBER: US/10/098,263B
/ CURRENT FILING DATE: 2003-01-08
/ PRIOR APPLICATION NUMBER: 60/276,759
/ PRIOR FILING DATE: 2001-03-16
/ NUMBER OF SEQ ID NOS: 131066
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 8462
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-098-263B-8462

Query Match          100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9
Db 13 ACAAATGTT 5

RESULT 10
US-10-098-263B-15021
/ Sequence 15021, Application US/10098263B
/ Publication No. US20030104410A1
/ GENERAL INFORMATION:
/ APPLICANT: Mittman, Michael
/ TITLE OF INVENTION: Human Microarray
/ FILE REFERENCE: 3118.1
/ CURRENT APPLICATION NUMBER: US/10/098,263B
/ CURRENT FILING DATE: 2003-01-08
/ PRIOR APPLICATION NUMBER: 60/276,759
/ PRIOR FILING DATE: 2001-03-16
/ NUMBER OF SEQ ID NOS: 131066
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 15021
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-098-263B-15021

Query Match          100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAATGTT 9
Db 3 ACAAATGTT 11

RESULT 11
US-10-098-263B-15022
```



```
; Sequence 15022, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 15022
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-15022
```

```
Query Match      100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ACAAATGTT 9
   |||||
Db 3 ACAAATGTT 11
```

## RESULT 12

```
US-10-098-263B-16053/c
; Sequence 16053, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 16053
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-16053
```

```
Query Match      100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ACAAATGTT 9
   |||||
Db 25 ACAAATGTT 17
```

## RESULT 13

```
US-10-098-263B-16054/c
; Sequence 16054, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 16054
; LENGTH: 25
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
US-10-098-263B-16054
```

```
Query Match      100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ACAAATGTT 9
   |||||
Db 25 ACAAATGTT 17
```

## RESULT 14

```
US-10-098-263B-22881
; Sequence 22881, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 22881
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-22881
```

```
Query Match      100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ACAAATGTT 9
   |||||
Db 16 ACAAATGTT 24
```

## RESULT 15

```
US-10-098-263B-22882
; Sequence 22882, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 22882
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-22882
```

```
Query Match      100.0%; Score 9; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ACAAATGTT 9
   |||||
Db 16 ACAAATGTT 24
```

```
Search completed: March 25, 2004, 15:52:48
Job time : 8.79622 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 07:40:23 ; Search time 9.33989 seconds  
(without alignments)  
4093.601 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_223\_231

Perfect score: 9  
Sequence: 1 acaaatgtt 9

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	100.0	15	6	ABN87945
C 2	9	100.0	17	2	AAx72750
C 3	9	100.0	17	2	AAx72751
C 4	9	100.0	17	2	AAV95125
C 5	9	100.0	17	2	AAV94640
C 6	9	100.0	17	2	AAV95126
C 7	9	100.0	17	2	AAV94639
C 8	9	100.0	17	3	AAf04437
C 9	9	100.0	17	3	AAf04884
C 10	9	100.0	17	3	AAf04436
C 11	9	100.0	17	3	AAf04885
C 12	9	100.0	17	7	AAf04885
C 13	9	100.0	17	7	AAf04885
C 14	9	100.0	17	9	ADb42327
C 15	9	100.0	17	9	ADb42327
C 16	9	100.0	18	2	AAf04437
C 17	9	100.0	18	2	AAf04437
C 18	9	100.0	18	6	ABL89242
C 19	9	100.0	18	6	ABL89256
C 20	9	100.0	18	6	ABL89240
C 21	9	100.0	18	6	ABL89251
C 22	9	100.0	18	6	ABL89237
C 23	9	100.0	19	3	AAa84073
C 24	9	100.0	19	3	AAa84071

24 9 100.0 19 3 AAa84072 Cyclin C  
25 9 100.0 19 3 AAa84070 Cyclin C  
26 9 100.0 19 3 AAH59233 Cyclin C  
27 9 100.0 19 5 AAH59232 Cyclin C  
28 9 100.0 19 5 AAH59235 Cyclin C  
29 9 100.0 19 5 AAH59234 Cyclin C  
30 9 100.0 20 2 AAZ04144 PCR prime  
31 9 100.0 20 3 AAa57975 Candida a  
32 9 100.0 20 3 ABLS3521 SAM1 gene  
33 9 100.0 20 6 ABa95214 Mouse SAM  
34 9 100.0 20 6 ABa95214 Mouse SAM  
35 9 100.0 20 6 ABa95214 Mouse SAM  
36 9 100.0 20 6 ABa95214 Mouse SAM  
37 9 100.0 20 7 ABZ93242 Human oli  
38 9 100.0 20 7 ABZ92680 Human oli  
39 9 100.0 20 7 ABZ92681 Human oli  
40 9 100.0 20 7 ABZ93243 Human oli  
41 9 100.0 20 7 ABT43215 Neuroblas  
42 9 100.0 20 7 ABT43215 Neuroblas  
43 9 100.0 20 9 ACf79545 SAM1 spec  
44 9 100.0 22 2 AAa41944 Nucleotid  
45 9 100.0 22 2 AAa37083 PCR prime

## ALIGNMENTS

RESULT 1  
ABN87945  
ID ABN87945 standard; DNA; 15 BP.  
XX  
AC ABN87945;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human GSR allele specific oligonucleotide primer SEQ ID NO:64.  
XX  
KW Human; glutathione reductase; GSR; enzyme; haemolytic anaemia; SNP;  
KW Gene therapy; antianaemic; polymorphic; single nucleotide polymorphism;  
KW primer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 14  
FT /tag= a  
FT /note= "polymorphic base"  
XX  
PN WO200242320-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 13-NOV-2001; 2001WO-US046473.  
XX  
PR 10-NOV-2000; 2000US-0247202P.  
XX  
PA (GENA-) GENAISSANCE PHARM INC.  
XX  
PI Bieglecki KM, Sanchis A, Sausker EA, Sun X;  
XX  
PS WPI; 2002-471719/50.  
XX  
PT New genetic variants of Glutathione reductase isogenes, useful for  
PT improving efficiency and reliability in drug development for treating  
PT hemolytic anemia.  
XX  
PS Claim 14; Page 14; 137pp; English.  
XX  
CC The present invention describes genetic variants of the human glutathione  
CC reductase (GSR) gene (I). (II) has antianaemic activity and can be used in  
CC gene therapy. (I) can be used in screening for drugs targeting (I) that  
CC are useful for treating haemolytic anaemia. Methods from the present  
CC invention can be used; for improving the efficiency and reliability of

CC several steps in the discovery and development of drugs for treating  
 CC diseases associated with GSR activity; for haplotyping, which is also  
 CC used by the pharmaceutical research scientist to validate GSR as a  
 CC candidate target for treating a specific condition or disease predicted  
 CC to be associated with GSR activity, e.g. haemolytic anaemia, and in the  
 CC design of clinical trials for treating a specific condition of disease  
 CC associated with GSR activity; and for screening compounds targeting GSR.  
 CC (1) is useful in studying the expression and function of GSR, and in  
 CC expressing GSR protein for use in screening for candidate drugs to treat  
 CC diseases related to GSR activity. (1) is also useful in studying the  
 CC effect of the variation on the biological activity of GSR as well as on  
 CC the binding affinity of candidate drugs targeting GSR for the treatment  
 CC of haemolytic anaemia. The present sequence represents an allele specific  
 CC oligonucleotide (ASO) primer for the human GSR gene, which is given in  
 CC the exemplification of the present invention. N.B. The polymorphic base  
 CC (showing a single nucleotide polymorphism) in the ASO primer is shown  
 CC using an IUPAC ambiguity code (as given in the present invention)  
 XX  
 SQ Sequence 15 BP; 4 A; 3 C; 2 G; 5 T; 0 U; 1 Other;

Query Match 100.0%; Score 9; DB 6; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
 |||||  
 Db 3 ACAAATGTT 11

## RESULT 2

AAAX72750/c  
 ID AAX72750 standard; RNA; 17 BP.  
 XX  
 AC AAX72750;  
 XX  
 DT 28-JUL-1999 (first entry)  
 XX  
 DE Mouse flk-1 VEGF receptor hammerhead ribozyme substrate #183.  
 XX  
 DE Vascular endothelial growth factor receptor; VEGF receptor; flt-1; flk-1;  
 KW KDR; hammerhead ribozyme; hairpin ribozyme; cleavage;  
 KW tumour angiogenesis; psoriasis; rheumatoid arthritis; ocular disease;  
 KW fms-like tyrosine kinase 1; kinase insert domain containing receptor;  
 KW foetal liver kinase 1; ss.  
 XX

OS Mus sp.

XX WO9715662-A2.

XX 01-MAY-1997.

XX 25-OCT-1996; 96WO-US017480.

XX 26-OCT-1995; 95US-0005974P.

XX 11-JAN-1996; 96US-00584040.

XX (RIBO-) RIBOZYME PHARM INC.

XX (CHIR) CHIRON CORP.

XX Pavco P, Mcswiggen J, Stinchcomb D, Escobedo J;

XX WPI; 1997-259017/23.

XX Nucleic acid molecule modulating VEGF receptor(s) gene expression or mRNA  
 PT stability - useful for treating e.g. tumour angiogenesis, psoriasis,  
 PT rheumatoid arthritis, etc., in a human patient.

XX Claim 4; Page 128; 218pp; English.

XX The present invention describes nucleic acid molecules which modulate the  
 CC synthesis, expression and/or stability of a mRNA encoding 1 or more  
 CC receptors of vascular endothelial growth factor (VEGF). A patient  
 CC (preferably human) having a condition associated with the level of the

CC fms-like tyrosine kinase 1 (flt-1), kinase insert domain containing  
 CC receptor (KDR) and/or foetal liver kinase 1 (flk-1) (e.g. tumour  
 CC angiogenesis, ocular diseases, psoriasis and rheumatoid arthritis) can be  
 CC treated by administering the nucleic acid molecule or the expression  
 CC vector to the patient. AAX7275 to AAX75752 represent specific examples  
 CC of nucleic acid molecules from the present invention  
 XX  
 SQ Sequence 17 BP; 5 A; 3 C; 4 G; 0 T; 5 U; 0 Other;

Query Match 100.0%; Score 9; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGTT 9  
 |||||  
 Db 12 ACAAATGTT 4

## RESULT 3

AAAX72751/c  
 ID AAX72751 standard; RNA; 17 BP.

XX  
 AC AAX72751;

XX 28-JUL-1999 (first entry)

XX Mouse flk-1 VEGF receptor hammerhead ribozyme substrate #184.

XX Vascular endothelial growth factor receptor; VEGF receptor; flt-1; flk-1;  
 KW KDR; hammerhead ribozyme; hairpin ribozyme; cleavage;  
 KW tumour angiogenesis; psoriasis; rheumatoid arthritis; ocular disease;  
 KW fms-like tyrosine kinase 1; kinase insert domain containing receptor;  
 KW foetal liver kinase 1; ss.

OS Mus sp.

XX WO9715662-A2.

XX 01-MAY-1997.

XX 25-OCT-1996; 96WO-US017480.

XX 26-OCT-1995; 95US-0005974P.

XX 11-JAN-1996; 96US-00584040.

XX (RIBO-) RIBOZYME PHARM INC.

XX (CHIR) CHIRON CORP.

XX Pavco P, Mcswiggen J, Stinchcomb D, Escobedo J;

XX WPI; 1997-259017/23.

XX Nucleic acid molecule modulating VEGF receptor(s) gene expression or mRNA  
 PT stability - useful for treating e.g. tumour angiogenesis, psoriasis,  
 PT rheumatoid arthritis, etc., in a human patient.

XX Claim 4; Page 128; 218pp; English.

XX The present invention describes nucleic acid molecules which modulate the  
 CC synthesis, expression and/or stability of a mRNA encoding 1 or more  
 CC receptors of vascular endothelial growth factor (VEGF). A patient  
 CC (preferably human) having a condition associated with the level of the  
 CC fms-like tyrosine kinase 1 (flt-1), kinase insert domain containing  
 CC receptor (KDR) and/or foetal liver kinase 1 (flk-1) (e.g. tumour  
 CC angiogenesis, ocular diseases, psoriasis and rheumatoid arthritis) can be  
 CC treated by administering the nucleic acid molecule or the expression  
 CC vector to the patient. AAX7275 to AAX75752 represent specific examples  
 CC of nucleic acid molecules from the present invention  
 XX

SQ Sequence 17 BP; 5 A; 3 C; 4 G; 0 T; 5 U; 0 Other;

Query Match 100.0%; Score 9; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+04;

```

Matches      9;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  ACAAATGTT 9
      |||||
Db      11  ACAAATGTT 3

RESULT 4
AAV95125/c
ID  AAV95125 standard; RNA; 17 BP.
XX
AC  AAV95125;
XX
DT  24-FEB-1999 (first entry)
XX
DE  Canine IL-2 receptor g-chain substrate position 383.
XX
KW  Human; IL-2 receptor g-chain; interleukin 2 receptor gamma chain;
KW  hammerhead ribozyme; hairpin ribozyme; substrate; expression; cancer;
KW  autoimmune disease; psoriasis; allergy; inflammatory disease;
KW  graft rejection; ss.
XX
OS  Synthetic.
OS  Canis sp.
XX
PN  WO9824913-A2.
XX
PD  11-JUN-1998.
XX
PF  02-DEC-1997; 97WO-US021748.
XX
PR  03-DEC-1996; 96US-00758306.
XX
PA  (RIBO-) RIBOZYME PHARM INC.
XX
PI  Stinchcomb DT, Mcswiggen JA;
XX
WPI; 1998-333332/29.
XX
PT  Ribozymes targeted to interleukin 2 - useful for treating e.g. cancer,
PT  autoimmune disease and allergies.
XX
PS  Claim 4; Page 46; 61pp; English.
XX
CC  The present sequence invention describes ribozymes targeted to modulate
CC  the synthesis and/or expression of interleukin (IL)-2R gamma encoded RNA.
CC  AAV93889 to AAV94574 represent specifically claimed ribozymes, and
CC  AAV94575 to AAV95260 represent specifically claimed substrate sequences
CC  from the present invention. The ribozymes can be used for the treatment
CC  of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy
CC  and other inflammatory conditions. The ribozymes are also used to induce
CC  tolerance in a recipient to alloantigen from a donor
XX
SQ  Sequence 17 BP; 4 A; 4 C; 3 G; 0 T; 6 U; 0 Other;

Query Match      100.0%; Score 9; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches      9;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

PS  Claim 4; Page 46; 61pp; English.
XX
CC  The present sequence invention describes ribozymes targeted to modulate
CC  the synthesis and/or expression of interleukin (IL)-2R gamma encoded RNA.
CC  AAV93889 to AAV94574 represent specifically claimed ribozymes, and
CC  AAV94575 to AAV95260 represent specifically claimed substrate sequences
CC  from the present invention. The ribozymes can be used for the treatment
CC  of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy
CC  and other inflammatory conditions. The ribozymes are also used to induce
CC  tolerance in a recipient to alloantigen from a donor
XX
SQ  Sequence 17 BP; 4 A; 4 C; 3 G; 0 T; 6 U; 0 Other;

Query Match      100.0%; Score 9; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches      9;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  ACAAATGTT 9
      |||||
Db      12  ACAAATGTT 4

RESULT 5
AAV94640/c
ID  AAV94640 standard; RNA; 17 BP.
XX
AC  AAV94640;
XX
DT  24-FEB-1999 (first entry)
XX
DE  Human IL-2 receptor g-chain substrate position 398.

```

```

XX
KW  Human; IL-2 receptor g-chain; interleukin 2 receptor gamma chain;
KW  hammerhead ribozyme; hairpin ribozyme; substrate; expression; cancer;
KW  autoimmune disease; psoriasis; allergy; inflammatory disease;
KW  graft rejection; ss.
XX
OS  Homo sapiens.
XX
PN  WO9824913-A2.
XX
PD  11-JUN-1998.
XX
PF  02-DEC-1997; 97WO-US021748.
XX
PR  03-DEC-1996; 96US-00758306.
XX
PA  (RIBO-) RIBOZYME PHARM INC.
XX
PI  Stinchcomb DT, Mcswiggen JA;
XX
WPI; 1998-333332/29.
XX
PT  Ribozymes targeted to interleukin 2 - useful for treating e.g. cancer,
PT  autoimmune disease and allergies.
XX
PS  Claim 4; Page 34; 61pp; English.
XX
CC  The present sequence invention describes ribozymes targeted to modulate
CC  the synthesis and/or expression of interleukin (IL)-2R gamma encoded RNA.
CC  AAV93889 to AAV94574 represent specifically claimed ribozymes, and
CC  AAV94575 to AAV95260 represent specifically claimed substrate sequences
CC  from the present invention. The ribozymes can be used for the treatment
CC  of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy
CC  and other inflammatory conditions. The ribozymes are also used to induce
CC  tolerance in a recipient to alloantigen from a donor
XX
SQ  Sequence 17 BP; 5 A; 3 C; 2 G; 0 T; 7 U; 0 Other;

Query Match      100.0%; Score 9; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches      9;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  ACAAATGTT 9
      |||||
Db      11  ACAAATGTT 3

RESULT 6
AAV95126/c
ID  AAV95126 standard; RNA; 17 BP.
XX
AC  AAV95126;
XX
DT  24-FEB-1999 (first entry)
XX
DE  Canine IL-2 receptor g-chain substrate position 384.
XX
KW  Human; IL-2 receptor g-chain; interleukin 2 receptor gamma chain;
KW  hammerhead ribozyme; hairpin ribozyme; substrate; expression; cancer;
KW  autoimmune disease; psoriasis; allergy; inflammatory disease;
KW  graft rejection; ss.
XX
OS  Synthetic.
OS  Canis sp.
XX
PN  WO9824913-A2.
XX
PD  11-JUN-1998.
XX
PF  02-DEC-1997; 97WO-US021748.
XX
PR  03-DEC-1996; 96US-00758306.

```

PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Stinchcomb DT, Mcswiggen JA;  
XX  
DR WPI; 1998-333332/29.  
XX  
PT Ribozymes targetted to interleukin 2 - useful for treating e.g. cancer,  
PT autoimmune disease and allergies.  
XX  
PS Claim 4; Page 46; 61pp; English.  
XX  
CC The present sequence invention describes ribozymes targeted to modulate  
CC the synthesis and/or expression of interleukin (IL)-2R gamma encoded RNA.  
CC AAV93889 to AAV94574 represent specifically claimed ribozymes, and  
CC AAV94575 to AAV95260 represent specifically claimed substrate sequences  
CC from the present invention. The ribozymes can be used for the treatment  
CC of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy  
CC and other inflammatory conditions. The ribozymes are also used to induce  
CC tolerance in a recipient to alloantigen from a donor  
XX  
SQ Sequence 17 BP; 5 A; 3 C; 3 G; 0 T; 6 U; 0 Other;  
Query Match 100.0%; Score 9; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAAAATGTT 9  
DB 11 ACAAAATGTT 3  
RESULT 7  
AAV94639/c  
ID AAV94639 standard; RNA; 17 BP.  
XX  
AC AAV94639;  
XX  
DT 24-FEB-1999 (first entry)  
XX  
DE Human IL-2 receptor g-chain substrate position 397.  
XX  
XX Human; IL-2 receptor g-chain; interleukin 2 receptor gamma chain;  
KW hammerhead ribozyme; hairpin ribozyme; substrate; expression; cancer;  
KW autoimmune disease; psoriasis; allergy; inflammatory disease;  
KW graft rejection; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9824913-A2.  
XX  
PD 11-JUN-1998.  
XX  
PF 02-DEC-1997; 97WO-US021748.  
XX  
PR 03-DEC-1996; 96US-00758306.  
XX  
PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Stinchcomb DT, Mcswiggen JA;  
XX  
DR WPI; 1998-333332/29.  
XX  
PT Ribozymes targetted to interleukin 2 - useful for treating e.g. cancer,  
PT autoimmune disease and allergies.  
XX  
PS Claim 4; Page 34; 61pp; English.  
XX  
CC The present sequence invention describes ribozymes targeted to modulate  
CC the synthesis and/or expression of interleukin (IL)-2R gamma encoded RNA.  
CC AAV93889 to AAV94574 represent specifically claimed ribozymes, and  
CC AAV94575 to AAV95260 represent specifically claimed substrate sequences  
CC from the present invention. The ribozymes can be used for the treatment  
CC of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy  
CC of, e.g. graft rejection, autoimmune disease, cancer, psoriasis, allergy  
XX  
CC and other inflammatory conditions. The ribozymes are also used to induce  
CC tolerance in a recipient to alloantigen from a donor  
XX  
SQ Sequence 17 BP; 5 A; 3 C; 3 G; 0 T; 6 U; 0 Other;  
Query Match 100.0%; Score 9; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAAAATGTT 9  
DB 11 ACAAAATGTT 3  
RESULT 8  
AAV94437/c  
ID AAV94437 standard; DNA; 17 BP.  
XX  
AC AAV94437;  
XX  
DT 16-FEB-2001 (first entry)  
XX  
DE Hammerhead ribozyme substrate #1953.  
XX  
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;  
KW interferon alpha; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200061729-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 11-APR-2000; 2000WO-US009721.  
XX  
PR 12-APR-1999; 99US-0129390P.  
XX  
PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Blatt L, Zwick M, Pavco P, Mcswiggen J;  
XX  
DR WPI; 2000-647423/62.  
XX  
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,  
PT useful for producing e.g. granulocyte colony stimulating factor protein,  
PT interferon alpha and erythropoietin.  
XX  
PS Claim 4; Page 100; 164pp; English.  
XX  
CC The present invention relates to enzymatic and antisense nucleic acid  
CC molecules that act as inhibitors of the expression of repressor genes  
CC encoding the TR2 Orphan receptor, EAR3/COUP-TF-1, the GATA transcription  
CC factor gene, IRF-2 and/or the CAAAT Displacement Protein (CDP).  
CC Inhibition of the repressors removes prevents inhibition (and  
CC consequently increases expression of) genes involved in the production of  
CC erythropoietin, granulocyte colony stimulating factor protein and  
CC interferon alpha  
XX  
SQ Sequence 17 BP; 5 A; 3 C; 3 G; 6 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAAAATGTT 9  
DB 11 ACAAAATGTT 3  
RESULT 9  
AAV94884/c  
ID AAV94884 standard; DNA; 17 BP.  
XX

AC AAF04884;  
XX  
DT 16-FEB-2001 (first entry)  
XX  
DE Hammerhead ribozyme substrate #2400.  
XX  
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;  
KW interferon alpha; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200061729-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 11-APR-2000; 2000WO-US009721.  
XX  
PR 12-APR-1999; 99US-0129390P.  
XX  
PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Blatt L, Zwick M, Pavco P, Mcswiggen J;  
XX  
DR WPI; 2000-647423/62.  
XX  
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,  
PT useful for producing e.g. granulocyte colony stimulating factor protein,  
PT interferon alpha and erythropoietin.  
XX  
PS Claim 4; Page 110; 164pp; English.  
XX  
CC The present invention relates to enzymatic and antisense nucleic acid  
CC molecules that act as inhibitors of the expression of repressor genes  
CC encoding the TR2 Orphan receptor, EAR3/COUP-TF-1, the GATA transcription  
CC factor gene, IRF-2 and/or the CAATT Displacement Protein (CDP).  
CC Inhibition of the repressors removes prevents inhibition (and  
CC consequently increases expression of) genes involved in the production of  
CC erythropoietin, granulocyte colony stimulating factor protein and  
CC interferon alpha  
XX  
SQ Sequence 17 BP; 6 A; 3 C; 2 G; 6 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAAAATGTT 9  
DB 12 ACAAAATGTT 4  
RESULT 10  
AAF04436/C  
ID AAF04436 standard; DNA; 17 BP.  
XX  
AC AAF04436;  
XX  
DT 16-FEB-2001 (first entry)  
XX  
DE Hammerhead ribozyme substrate #1952.  
XX  
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;  
KW interferon alpha; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200061729-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 11-APR-2000; 2000WO-US009721.  
XX  
PR 12-APR-1999; 99US-0129390P.  
XX

XX (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Blatt L, Zwick M, Pavco P, Mcswiggen J;  
XX  
DR WPI; 2000-647423/62.  
XX  
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,  
PT useful for producing e.g. granulocyte colony stimulating factor protein,  
PT interferon alpha and erythropoietin.  
XX  
PS Claim 4; Page 100; 164pp; English.  
XX  
CC The present invention relates to enzymatic and antisense nucleic acid  
CC molecules that act as inhibitors of the expression of repressor genes  
CC encoding the TR2 Orphan receptor, EAR3/COUP-TF-1, the GATA transcription  
CC factor gene, IRF-2 and/or the CAATT Displacement Protein (CDP).  
CC Inhibition of the repressors removes prevents inhibition (and  
CC consequently increases expression of) genes involved in the production of  
CC erythropoietin, granulocyte colony stimulating factor protein and  
CC interferon alpha  
XX  
SQ Sequence 17 BP; 6 A; 3 C; 2 G; 6 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAAAATGTT 9  
DB 12 ACAAAATGTT 4  
RESULT 11  
AAF04885/C  
ID AAF04885 standard; DNA; 17 BP.  
XX  
AC AAF04885;  
XX  
DT 16-FEB-2001 (first entry)  
XX  
DE Hammerhead ribozyme substrate #2401.  
XX  
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;  
KW interferon alpha; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200061729-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 11-APR-2000; 2000WO-US009721.  
XX  
PR 12-APR-1999; 99US-0129390P.  
XX  
PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Blatt L, Zwick M, Pavco P, Mcswiggen J;  
XX  
DR WPI; 2000-647423/62.  
XX  
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,  
PT useful for producing e.g. granulocyte colony stimulating factor protein,  
PT interferon alpha and erythropoietin.  
XX  
PS Claim 4; Page 110; 164pp; English.  
XX  
CC The present invention relates to enzymatic and antisense nucleic acid  
CC molecules that act as inhibitors of the expression of repressor genes  
CC encoding the TR2 Orphan receptor, EAR3/COUP-TF-1, the GATA transcription  
CC factor gene, IRF-2 and/or the CAATT Displacement Protein (CDP).  
CC Inhibition of the repressors removes prevents inhibition (and  
CC consequently increases expression of) genes involved in the production of  
CC erythropoietin, granulocyte colony stimulating factor protein and  
CC interferon alpha  
XX

```
CC consequently increases expression of) genes involved in the production of
CC erythropoietin; granulocyte colony stimulating factor protein and
CC interferon alpha
XX
SQ Sequence 17 BP; 5 A; 3 C; 3 G; 6 T; 0 U; 0 Other;
  Query Match 100.0%; Score 9; DB 3; Length 17;
  Best Local Similarity 100.0%; Pred. No. 3.2e+04;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAAAATGTT 9
Db 11 ACAAAATGTT 3
  |||||
  |||||

RESULT 12
ACCS4403
ID ACC54403 standard; DNA; 17 BP.
XX
AC ACC54403;
XX
27-JUN-2003 (first entry)
XX
DE Human tumour suppressor sequence #3170.
XX
ss; tumour suppressor; antitumour; cytostatic; tumour suppression;
KW tumour regression; apoptosis; virus resistance; diagnosis;
KW cellular degeneration.
XX
OS Homo sapiens.
XX
PN FR2826373-A1.
XX
PD 27-DEC-2002.
XX
20-JUN-2001; 2001PR-00008139.
XX
20-JUN-2001; 2001PR-00008139.
XX
PA (MOLE-) MOLECULAR ENGINES LAB SA.
XX
PI Tuijnder M, Telerman A, Anson R;
XX
WPI; 2003-250498/25.
XX
New nucleic acid sequences associated with tumor suppression, regression,
PT apoptosis or virus resistance are useful to diagnose and treat viral
PT disease, development of tumor cells and cell degeneration.
XX
Claim 1; Page 772; 798pp; French.
XX
This sequence represents an isolated nucleic acid sequence associated
CC with tumour suppression or regression, apoptosis or virus resistance. The
CC invention relates to these sequences or sequences having at least 80%
CC identity to them, and polypeptides encoded by the sequences or
CC polypeptides having 80% identity to the polypeptide sequences. The
CC invention is used to diagnose or treat viral disease or disease
CC characterized by development of tumour cells or cellular degeneration
XX
Sequence 17 BP; 6 A; 2 C; 2 G; 7 T; 0 U; 0 Other;
  Query Match 100.0%; Score 9; DB 7; Length 17;
  Best Local Similarity 100.0%; Pred. No. 3.2e+04;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAAAATGTT 9
Db 9 ACAAAATGTT 17
  |||||
  |||||

RESULT 13
ADB42327
ID ADB42327 standard; DNA; 17 BP.
```

```
XX ADB42327;
AC
XX 18-DEC-2003 (revised)
DT 04-DEC-2003 (first entry)
XX
XX Tumour suppression/reversion associated nucleotide #2650.
DE
XX cytostatic; antiviral; neuroprotective; nootropic; neuroleptic; ss;
KW primer; probe; tumour suppression; tumour reversion; apoptosis;
KW virus resistance; transgenic animals; Alzheimer's disease; schizophrenia;
KW diagnosis.
XX
XX Homo sapiens.
XX
PN WO2003040369-A2.
XX
15-MAY-2003.
XX
17-SEP-2002; 2002WO-IB004219.
XX
17-SEP-2001; 2001PR-00011981.
XX
(MOLE-) MOLECULAR ENGINES LAB.
XX
Telerman A, Anson R, Tuijnder M;
XX
WPI; 2003-441574/41.
XX
New nucleic acid encoding human prostate membrane-specific antigen,
PT useful e.g. for treatment of tumors and viral infection, also related
PT polypeptide and antibodies.
XX
Disclosure; Page 341; 771pp; French.
XX
The invention relates to the isolation of 6327 nucleotide sequences,
CC fragments of at least 15 consecutive nucleotides of these nucleotides, a
CC sequence having at least 80% identity, after optimal alignment, with the
CC nucleotides, a sequence that hybridizes under stringent conditions with
CC the nucleotides, or the complement or corresponding RNA, of the
CC nucleotides. The nucleotides are used as probes or primers for detecting,
CC identifying, quantifying and/or amplifying nucleic acids, as in vitro
CC sense and antisense sequences, of nucleotides involved in tumour
CC recombinant polypeptides, and to prepare transgenic animals, as
CC experimental models. The nucleotides (also vectors containing them and
CC cells containing the vectors), the encoded polypeptides and antibodies
CC (Ab), against the polypeptide are useful for prevention and/or treatment
CC of viral infections or diseases characterized by development of tumours
CC or cell degeneration (e.g. Alzheimer's disease or schizophrenia).
CC Analysis of the expression of the nucleotides can be used for diagnosis
CC and/or prognosis of these diseases. The nucleotides and polypeptides can
CC also be used to screen for their specific interactive molecules,
CC potentially useful for treating diseases associated with abnormal
CC expression of the nucleotides.
SQ Sequence 17 BP; 6 A; 2 C; 2 G; 7 T; 0 U; 0 Other;
  Query Match 100.0%; Score 9; DB 9; Length 17;
  Best Local Similarity 100.0%; Pred. No. 3.2e+04;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAAAATGTT 9
Db 9 ACAAAATGTT 17
  |||||
  |||||

RESULT 14
ADB44770
ID ADB44770 standard; DNA; 17 BP.
XX
AC ADB44770;
XX
```

DT 18-DEC-2003 (first entry)  
 DE Tumour suppression/reversion associated nucleotide #5093.  
 DE cytostatic; antiviral; neuroprotective; neurotropic; neuroleptic; ss;  
 KW primer; probe; tumour suppression; tumour reversion; apoptosis;  
 KW virus resistance; transgenic animals; Alzheimer's disease; schizophrenia;  
 KW diagnosis.  
 XX Homo sapiens.  
 OS  
 XX WO2003040369-A2.  
 PN  
 XX 15-MAY-2003.  
 PD  
 XX 17-SEP-2002; 2002WO-IB004219.  
 PF  
 XX 17-SEP-2001; 2001FR-00011981.  
 PR  
 XX (MOLE-) MOLECULAR ENGINES LAB.  
 PA  
 XX Tellerman A, Amson R, Tuijnder M;  
 PI  
 XX WPI; 2003-441574/41.  
 DR  
 XX New nucleic acid encoding human prostate membrane-specific antigen,  
 PT useful e.g. for treatment of tumors and viral infection, also related  
 PT polypeptide and antibodies.  
 PT  
 PS Disclosure; Page 627; 771pp; French.  
 XX  
 CC The invention relates to the isolation of 6327 nucleotide sequences,  
 CC fragments of at least 15 consecutive nucleotides of these nucleotides, a  
 CC sequence having at least 80% identity, after optimal alignment, with the  
 CC nucleotides, a sequence that hybridizes under stringent conditions with  
 CC the nucleotides, or the complement, or corresponding RNA, of the  
 CC nucleotides. The nucleotides are used as probes or primers for detecting,  
 CC identifying, quantifying and/or amplifying nucleic acids, as in vitro  
 CC sense and antisense sequences, of nucleotides involved in tumour  
 CC suppression or reversion, apoptosis and or viral resistance, to produce  
 CC recombinant polypeptides, and to prepare transgenic animals, as  
 CC experimental models. The nucleotides (also vectors containing them and  
 CC cells containing the vectors), the encoded polypeptides and antibodies  
 CC (Ab) against the polypeptide are useful for prevention and/or treatment  
 CC of viral infections or diseases characterized by development of tumours  
 CC or cell degeneration (e.g. Alzheimer's disease or schizophrenia).  
 CC Analysis of the expression of the nucleotides can be used for diagnosis  
 CC and/or prognosis of these diseases. The nucleotides and polypeptides can  
 CC also be used to screen for their specific interactive molecules,  
 CC potentially useful for treating diseases associated with abnormal  
 CC expression of the nucleotides.  
 XX  
 SQ Sequence 17 BP; 6 A; 3 C; 3 G; 5 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 9; DB 9; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACAAAATGTT 9  
 Db |||||  
 5 ACAAAATGTT 13  
 RESULT 15  
 AAT91799  
 ID AAT91799 standard; DNA; 18 BP.  
 XX  
 AC AAT91799;  
 XX  
 DT 14-APR-1998 (first entry)  
 XX  
 DE Dihydropyrimidine dehydrogenase genomic fragment PCR primer DPD15.  
 XX

KW Human; dihydropyrimidine dehydrogenase; DPD; slicing defect; detection;  
 KW 5-fluorouracil; cancer; anticancer; uraciluria; PCR primer; ss.  
 XX Synthetic.  
 OS  
 XX WO9735034-A1.  
 PN  
 XX 25-SEP-1997.  
 PD  
 XX 19-MAR-1997; 97WO-US004269.  
 PF  
 XX 20-MAR-1996; 96US-0013835P.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Gonzalez FU, Fernandez-Salguero P;  
 PI  
 XX WPI; 1997-480236/44.  
 DR  
 XX Detecting a splicing defect in the dihydro:pyrimidine dehydrogenase gene  
 PT - used to identify subjects sensitive to 5-fluorouracil, toxic to  
 PT individuals with DPD defects.  
 PT  
 XX Example 2; Page 24; 38pp; English.  
 PS  
 XX A novel method has been developed for detecting the splicing defect in  
 CC the dihydropyrimidine dehydrogenase (DPD) gene. The method comprises  
 CC determining whether genomic DNA containing the DPD gene has a wild-type  
 CC intron-exon boundary for an exon that encodes amino acids (aa) 581-635 of  
 CC the corresponding DPD protein. The present sequence represents a PCR  
 CC primer for synthesising a probe from the DPD gene, used in an example of  
 CC the present invention. The method is used specifically to determine  
 CC sensitivity of subjects to the anticancer agent 5-fluorouracil, which is  
 CC dangerously toxic to those with DPD defects. It can also be used to  
 CC diagnose DPD-deficiency disorders such as uraciluria  
 XX  
 SQ Sequence 18 BP; 4 A; 4 C; 3 G; 7 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 9; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACAAAATGTT 9  
 Db |||||  
 6 ACAAAATGTT 14  
 Search completed: March 25, 2004, 10:25:12  
 Job time : 12.3399 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 08:39:03 ; Search time 115.83 Seconds  
(without alignments)  
6361.316 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375

Perfect score: 17

Sequence: 1 tgtggaagaataaata 17

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.in.\*

32: em.htg.other.\*

33: em.htg.pln.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.man.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	17	100.0	799	11	BV075239	BV075239 S212P6011
C 2	17	100.0	39961	9	AC000044	AC000044 Homo sapi
C 3	17	100.0	43738	9	AC000034	AC000034 Homo sapi
C 4	17	100.0	44532	9	AC109830	AC109830 Homo sapi
C 5	17	100.0	80543	2	AC090287	AC090287 Homo sapi
C 6	17	100.0	100116	2	AC074171	AC074171 Mus muscu
C 7	17	100.0	130604	9	HS1007G16	HS1007G16 Human DNA
C 8	17	100.0	135331	9	AC008798	AC008798 Homo sapi
C 9	17	100.0	158888	9	AC079112	AC079112 Homo sapi
C 10	17	100.0	156165	2	AC021453	AC021453 Homo sapi
C 11	17	100.0	157980	9	AC100852	AC100852 Homo sapi
C 12	17	100.0	158541	9	AC024563	AC024563 Homo sapi
C 13	17	100.0	168656	9	AC009108	AC009108 Homo sapi
C 14	17	100.0	169226	2	AC107725	AC107725 Mus muscu
C 15	17	100.0	170749	9	AC025272	AC025272 Homo sapi
C 16	17	100.0	173965	9	AC018890	AC018890 Homo sapi
C 17	17	100.0	178405	2	AC025927	AC025927 Homo sapi
C 18	17	100.0	184917	2	AC073732	AC073732 Mus muscu
C 19	17	100.0	187109	2	AC074100	AC074100 Homo sapi
C 20	17	100.0	187888	9	AC011467	AC011467 Homo sapi
C 21	17	100.0	192667	2	AC068140	AC068140 Mus muscu
C 22	17	100.0	195414	2	AC073696	AC073696 Mus muscu
C 23	17	100.0	195414	2	AC073696	AC073696 Mus muscu
C 24	17	100.0	207235	2	AC044805	AC044805 Mus muscu
C 25	17	100.0	233228	2	AC098171	AC098171 Rattus no
C 26	17	100.0	235090	2	AC103148	AC103148 Rattus no
C 27	17	100.0	239426	2	AC134476	AC134476 Rattus no
C 28	17	100.0	244236	2	AC144505	AC144505 Pan trogl
C 29	17	100.0	244236	2	AC144505	AC144505 Pan trogl
C 30	17	100.0	254292	2	AC133271	AC133271 Rattus no
C 31	17	100.0	307511	2	AC119644	AC119644 Rattus no
C 32	16	94.1	1807	6	E13051	E13051 cDNA encodi
C 33	16	94.1	4100	5	XLU13674	XLU13674 Xenopus lae
C 34	16	94.1	5334	9	AF330558	AF330558 Homo sapi
C 35	16	94.1	6876	9	AF335591	AF335591 Homo sapi
C 36	16	94.1	17429	9	AL445253	AL445253 Human DNA
C 37	16	94.1	65750	9	AC005003	AC005003 Homo sapi
C 38	16	94.1	67172	2	AC126364	AC126364 Homo sapi
C 39	16	94.1	76999	9	AP000795	AP000795 Homo sapi
C 40	16	94.1	91010	2	AC008811	AC008811 Homo sapi
C 41	16	94.1	93177	9	AC098678	AC098678 Homo sapi
C 42	16	94.1	101033	9	AC104090	AC104090 Homo sapi
C 43	16	94.1	104630	9	AC007848	AC007848 Homo sapi
C 44	16	94.1	106689	9	AC008919	AC008919 Homo sapi
C 45	16	94.1	107706	9	AC138123	AC138123 Homo sapi

ALIGNMENTS

RESULT 1  
BV075239/c  
LOCUS BV075239  
DEFINITION S212P6011.FG3.TO CZECHII/E1 Mus musculus STS genomic, sequence  
tagged site.  
ACCESSION BV075239  
VERSION BV075239.1 GI:31191034  
KEYWORDS STS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 799)  
AUTHORS Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,  
Lander, E.S., Lindblad-Toh, K. and Daly, M.J.

**TITLE** The mosaic structure of variation in the laboratory mouse genome  
**JOURNAL** Nature 420 (6915), 574-578 (2002)  
**MEDLINE** 22354684  
**PUBMED** 12466852  
**COMMENT**

Contact: Kerstin Lindblad-Toh  
 Whitehead Institute for Biomedical Research, Center for Genome  
 Research  
 320 Charles Street, Cambridge, MA 02141, USA  
 Tel: 6172521477  
 Fax: 6172580903  
 Email: kersli@genome.wi.mit.edu  
 Primer A: None  
 Primer B: None  
 STS size: 799  
 Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
 were generated from 12981/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS  
 reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP  
 detection was carried out by SSAHA-SNP. 225,000 reads were  
 annotated  
 as STSS and 81,000 SNPs were annotated with alleles from C57BL/6J  
 and the strain from which the particular read came. The validation  
 rate for these SNPs was estimated at approximately 98%.

**FEATURES** Location/Qualifiers  
 1..799  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /map="4 16 27-755 66714317-66715044"  
 /clone\_lib="CZCH11/EI"  
 <1..799

**ORIGIN**  
 Query Match 100.0%; Score 17; DB 11; Length 799;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-02;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAGGAAATAAATA 17  
 |||||  
 DB 25 TGTGGAGGAAATAAATA 9

**RESULT 2**  
**AC000044**  
**LOCUS** Homo sapiens Chromosome 22q13 Cosmid Clone p76e10, complete  
**DEFINITION** sequence.  
**AC000044**  
**VERSION** AC000044.3 GI:22475335  
**KEYWORDS** HTG.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 39961)  
 Fransson, I. and Dumanaki, J. P.  
**TITLE** A cosmid clone in meningioma deletion region of 22q13  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 39961)  
**AUTHORS** Pan, H., Lin, S. and Roe, B. A.  
**TITLE** A cosmid clone in meningioma deletion region of 22q13  
**JOURNAL** Unpublished  
**REFERENCE** 3 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (21-OCT-1996) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REMARK** p76e10  
**REFERENCE** 4 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.

**TITLE** Direct Submission  
 Submitted (09-SEP-1997) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REMARK** p76e10  
**REFERENCE** 5 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (18-FEB-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 6 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (11-MAR-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 7 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (21-MAR-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 8 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (17-APR-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 9 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (11-JUN-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 10 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (01-AUG-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 11 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (20-AUG-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 12 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (01-OCT-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 13 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (14-OCT-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 14 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (30-OCT-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 15 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.  
**TITLE** Direct Submission  
 Submitted (05-NOV-1998) Department Of Chemistry and Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
**REFERENCE** 16 (bases 1 to 39961)  
**AUTHORS** Roe, B. A.

TITLE  
JOURNAL  
Submitted (13-NOV-1998) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
17 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (16-JAN-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
18 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (19-JAN-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
19 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (21-JAN-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
20 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (22-JAN-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
21 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (12-APR-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
22 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (13-APR-1999) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
23 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (01-JUN-2000) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
24 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (25-AUG-2002) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
25 (bases 1 to 39961)  
Roe, B.A.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (09-APR-2003) Department Of Chemistry and Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
On Aug 25, 2002 this sequence version replaced gi:4581185.  
Because these overlapping clones came from different libraries.  
FEATURES  
source  
1..39961  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="22q13"  
/clone="p76e10"

Query Match 100.0%; Score 17; DB 9; Length 39961;  
Best Local Similarity 100.0%; Pred. NO. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGAATAAATA 17  
|||||  
Db 37916 TGTGGAAGGAATAAATA 37932  
RESULT 3  
AC000034 43738 bp DNA linear PRI 09-APR-2003  
LOCUS Homo sapiens Chromosome 22q13 Cosmid Clone 111194 In Meningioma  
DEFINITION Deletion Region, complete sequence.  
ACCESSION AC000034  
VERSION AC000034  
KEYWORDS HTG. AC000034.3 GI:22597444  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Fransson, I. and Dumanski, J.P.  
JOURNAL Unpublished  
TITLE 1 (bases 1 to 43738)  
REFERENCE 2 (bases 1 to 43738)  
AUTHORS Pan, H., Lin, S. and Roe, B.A.  
JOURNAL Unpublished  
TITLE Homo sapiens Cosmid Clone 111194 In Meningioma Deletion Region  
REFERENCE 3 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (21-OCT-1996) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REMARK 111194  
REFERENCE 4 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (09-SEP-1997) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REMARK 111194  
REFERENCE 5 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (18-FEB-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 6 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (28-MAR-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 7 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (25-AUG-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 8 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (01-OCT-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 9 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (14-OCT-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 10 (bases 1 to 43738)  
AUTHORS Roe, B.A.  
JOURNAL Direct Submission  
TITLE Submitted (30-OCT-1998) Department Of Chemistry And Biochemistry,



McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC015631, 2000 bp overlap; the clone sequenced to the right is RPI1-400D2, 2000 bp overlap. Actual end is at base position 85524 of RPI1-400D2.

#### FEATURES

```

Source
1..44532
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RPI1-122688"
/clone_lib="RPI-11"
63..142
/rpt_family="L2"
repeat_region
697..731
/rpt_family="AT_rich"
repeat_region
1274..1319
/rpt_family="(TCTA)n"
repeat_region
1428..1478
/rpt_family="AT_rich"
repeat_region
2447..2850
/rpt_family="ERV1"
repeat_region
2915..3149
/rpt_family="MIR"
repeat_region
3643..4062
/rpt_family="MaLR"
repeat_region
4418..4648
/rpt_family="MIR"
repeat_region
5322..5350
/rpt_family="AT_rich"
repeat_region
5400..5483
/rpt_family="MIR"
repeat_region
5466..5853
/rpt_family="L2"
repeat_region
5884..6983
/rpt_family="L1"
repeat_region
7909..7936
/rpt_family="(TTTA)n"
repeat_region
7939..8944
/rpt_family="L1"
repeat_region
9538..9572
/rpt_family="AT_rich"
repeat_region
11519..11638
/rpt_family="MIR"
repeat_region
11820..12393
/rpt_family="L1"
repeat_region
12686..12756
/rpt_family="Alu"
repeat_region
12758..12888
/rpt_family="L1"
repeat_region
13182..13413
/rpt_family="L1"
repeat_region
18036..18137
/rpt_family="AT_rich"
repeat_region
18540..18883
/rpt_family="ERV1"
repeat_region
18954..19004
/rpt_family="(CATATA)n"
19083..19203
/rpt_family="MER1_type"
repeat_region
19247..19288
/rpt_family="(CA)n"
repeat_region
19455..19488
/rpt_family="AT_rich"
repeat_region
19709..19755
/rpt_family="AT_rich"
repeat_region
20114..20167
/rpt_family="L2"
repeat_region
20467..20833
/rpt_family="MaLR"
repeat_region
20876..21032
/rpt_family="L2"
repeat_region
21612..21887
/rpt_family="Alu"
repeat_region
21893..22075
/rpt_family="L1"
repeat_region
22435..23791
/rpt_family="L1"
repeat_region
23792..24476
/rpt_family="L1"
misc_feature
24594..24960
/note="match to EST BI45889 (NID:G15249545)"
repeat_region
24660..24772
/rpt_family="MIR"
repeat_region
25939..26066
/rpt_family="MIR"
repeat_region
26474..26975
/rpt_family="MaLR"
misc_feature
26615..26806
/note="match to EST BI45889 (NID:G15249545)"
repeat_region
27142..27187
/rpt_family="AT_rich"
repeat_region
27569..27894
/rpt_family="Alu"
repeat_region
29779..29801
/rpt_family="(TA)n"
repeat_region
29802..30127
/rpt_family="MER2_type"
repeat_region
30128..30184
/rpt_family="(TA)n"
repeat_region
30716..31118
/rpt_family="ERV1"
repeat_region
31162..31216
/rpt_family="ERV1"
repeat_region
31353..31388
/rpt_family="(TA)n"
repeat_region
31513..31702
/rpt_family="MER1_type"
repeat_region
31735..31758
/rpt_family="AT_rich"
repeat_region
31768..32055
/rpt_family="Alu"
repeat_region
32493..32513
/rpt_family="(TA)n"
repeat_region
33000..33140
/rpt_family="MIR"
repeat_region
33201..33512
/rpt_family="Alu"

```

Query Match 100.0%; Score 17; DB 9; Length 44532;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0;

Qy 1 TGTGGAAGGAATAAATA 17

|||||

Db 41623 TGTGGAAGGAATAAATA 41607

RESULT 5  
 AC090287  
 LOCUS

AC090287 90543 bp DNA linear PRI 28-DEC-2001



```

repeat_region 12110..12133
/rpt_family="ATTGn"
repeat_region complement(12134..12212)
/rpt_family="FLAM"
repeat_region complement(12482..12557)
/rpt_family="MER103"
repeat_region complement(12558..12820)
/rpt_family="AluY"
repeat_region complement(12821..12924)
/rpt_family="MER103"
repeat_region 14326..14369
/rpt_family="AT_rich"
repeat_region 14689..15341
/rpt_family="L2"
repeat_region 15507..15734
/rpt_family="MLTIL"
repeat_region 15735..16044
/rpt_family="AluYb"
unsure complement(15892..15899)
/note="<30 qual SNGL region."
repeat_region 16045..16334
/rpt_family="AluSp"
repeat_region 16335..16377
/rpt_family="MLTIL"
repeat_region complement(16378..16869)
/rpt_family="Charlie4a"
repeat_region 16870..17112
/rpt_family="MLTIL"
repeat_region complement(17591..17740)
/rpt_family="MERSA"
unsure 17745..17750
/note="<30 qual SNGL region."
repeat_region 17837..17887
/rpt_family="AluYb8"
repeat_region 17874..17939
/rpt_family="FLAM"
repeat_region complement(18030..18868)
/rpt_family="MER63"
unsure 18236..18242
/note="<30 qual SNGL region."
unsure 18391..18395
/note="<30 qual SNGL region."
repeat_region complement(18869..19290)
/rpt_family="MLTIA2"
unsure 18965..18977
/note="<30 qual SNGL region."
repeat_region complement(19291..19390)
/rpt_family="MER63"
repeat_region complement(20076..20167)
/rpt_family="MIR3"
repeat_region 20494..20572
/rpt_family="MIR"
repeat_region 20605..20915

Query Match 100.0%; Score 17; DB 9; Length 90543;
Best Local Similarity 100.0%; Pred.No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGATAAATA 17
DB 83407 TGTGGAAGGATAAATA 83423

RESULT 6
AC074171/c
LOCUS Mus musculus clone RP23-77J9, WORKING DRAFT SEQUENCE, 25 unordered
DEFINITION pieces.
ACCESSION AC074171
VERSION AC074171.1 GI:9211281
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 100116)  
DOE Joint Genome Institute.  
TITLE Sequencing of Mouse  
JOURNAL Unpublished  
2 (bases 1 to 100116)  
DOE Joint Genome Institute.  
DIRECT SUBMISSION  
Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1766665  
Center clone name: RPCI-23\_77J9  
-----  
Summary Statistics  
Consensus quality: 83951 bases at least Q40  
Consensus quality: 91885 bases at least Q30  
Consensus quality: 93406 bases at least Q20  
Estimated insert size: 128000; agarose-fp estimation  
Estimated insert size: 97716; sum-of-contigs estimation  
Quality coverage: 10.11 in Q20 bases; agarose-fp estimation  
Quality coverage: 13.24 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1295: contig of 1295 bp in length  
\* 1296 1395: gap of unknown length  
\* 1396 2472: contig of 1077 bp in length  
\* 2473 2572: gap of unknown length  
\* 2573 3670: contig of 1098 bp in length  
\* 3671 3770: gap of unknown length  
\* 3771 5091: contig of 1321 bp in length  
\* 5092 5191: gap of unknown length  
\* 5192 7124 7223: contig of 1932 bp in length  
\* 7224 8307: contig of 1084 bp in length  
\* 8308 9914: contig of 1507 bp in length  
\* 9915 10014: gap of unknown length  
\* 10015 11274: contig of 1260 bp in length  
\* 11275 11374: gap of unknown length  
\* 11375 13332: contig of 1958 bp in length  
\* 13333 13432: gap of unknown length  
\* 13433 14503: contig of 1071 bp in length  
\* 14504 14603: gap of unknown length  
\* 14604 19151: contig of 4548 bp in length  
\* 19152 23580: contig of 4329 bp in length  
\* 23581 23680: gap of unknown length  
\* 23681 26374: contig of 2694 bp in length  
\* 26375 26474: gap of unknown length  
\* 26475 30780: contig of 4306 bp in length  
\* 30781 34234: contig of 3354 bp in length  
\* 34235 34334: gap of unknown length  
\* 34335 38616: contig of 4282 bp in length  
\* 38617 38716: gap of unknown length  
\* 38717 43921: contig of 5105 bp in length  
\* 43922 49110: contig of 5189 bp in length  
\* 49111 55314: contig of 6104 bp in length  
\* 49211





Sw:Q33963 Sw:Q05045 Tr:O97480 Sw:Q95058 Tr:O97131		/note="MIR repeat: matches 3. .118 of consensus"
Sw:P79278 Sw:Q02CT7 Sw:Q82574 Sw:P49464 Tr:Q51919	repeat_region	38392. .38775
Sw:P69578 Sw:Q69289 Sw:P08927 Sw:Q94596 Tr:Q60024	repeat_region	/note="L2 repeat: matches 2288. .2675 of consensus"
Tr:P90622 Sw:Q34198 Sw:P51349 Sw:P37282 Sw:P235861	repeat_region	38796. .39003
Sw:P35862 Sw:Q05972 Tr:P93370	repeat_region	/note="MERS8A repeat: matches 1. .223 of consensus"
/pseudo	repeat_region	39262. .39311
/codon_start=1	repeat_region	/note="L2 repeat: matches 2699. .2748 of consensus"
/evidence=not experimental	repeat_region	40069. .40340
complement(24669. .25200)	repeat_region	/note="L2 repeat: matches 2386. .2693 of consensus"
/note="match: GSS: Em:AQ504827"	repeat_region	41945. .42164
complement(25663. .26219)	repeat_region	/note="MIR repeat: matches 12. .259 of consensus"
/note="match: GSS: Em:AQ309682"	repeat_region	42813. .42903
26085. .26326	repeat_region	/note="MIR repeat: matches 48. .137 of consensus"
/note="match: GSS: Em:AQ463218"		
26086. .26569		
/note="match: GSS: Em:AQ347313"		
26106. .26574		
/note="match: GSS: Em:AQ405763"		
26294. .26464		
/note="L2 repeat: matches 2362. .2529 of consensus"		
27969. .28555		
/note="L2 repeat: matches 1685. .2333 of consensus"		
28558. .28868		
/note="MERS8B repeat: matches 23. .337 of consensus"		
29639. .29709		
/note="LTR33 repeat: matches 44. .119 of consensus"		
29715. .30122		
/note="LTR33 repeat: matches 102. .517 of consensus"		
30214. .30320		
/note="MER20 repeat: matches 2. .102 of consensus"		
30357. .30478		
/note="MER20 repeat: matches 98. .218 of consensus"		
30592. .30802		
/note="MIR repeat: matches 1. .216 of consensus"		
31220. .31393		
/note="MIR repeat: matches 56. .226 of consensus"		
32435. .32652		
/note="MIR repeat: matches 2. .261 of consensus"		
complement(join(32897. .33009,104107. .104314))		
/gene="dJ1007G16.2"		
complement(<32897. .>.33009)		
/gene="dJ1007G16.2"		
/note="exon borders derived from GENESCAN predictions; true borders could be further down- and upstream		
this gene fragment and fragment .2 (and possibly fragmented		
.1) are part of the same gene with a yet to be determined		
exon structure		
match: Proteins: Tr:O60494 Tr:Q9TU53 Sw:P13497 Tr:Q9T135		
Tr:Q43897 Tr:Q9Y561 Sw:P98069 Tr:Q9U000 Sw:Q15113		
Sw:Q61398 Tr:O4072 Tr:Q9Y6L7 Tr:O57460 Tr:O70244		
Tr:Q57381 Tr:O57382 Tr:Q9WYV6 Sw:O08628 Tr:Q91925		
Tr:Q62381		
/codon_start=2		
/evidence=not experimental		
/product="dJ1007G16.3 (novel CUB domain protein)"		
/protein_id="CAC05319.1"		
/db_xref="GI:12314082"		
/translation="FTCASLPAPVISKNWLRLHFTSDGNHRQRGFSAQVQ"		
33444. .33622		
/note="MIR repeat: matches 29. .251 of consensus"		
33764. .33933		
/note="MIR repeat: matches 76. .247 of consensus"		
34160. .34582		
/note="LTR33 repeat: matches 1. .450 of consensus"		
35326. .35657		
/note="L1ME1 repeat: matches 5804. .6135 of consensus"		
35660. .35774		
/note="FLAM_C repeat: matches 1. .115 of consensus"		
35903. .36201		
/note="L2 repeat: matches 2178. .2488 of consensus"		
36373. .36486		
Sw:Q33963 Sw:Q05045 Tr:O97480 Sw:Q95058 Tr:O97131		
Sw:P79278 Sw:Q02CT7 Sw:Q82574 Sw:P49464 Tr:Q51919		
Sw:P69578 Sw:Q69289 Sw:P08927 Sw:Q94596 Tr:Q60024		
Tr:P90622 Sw:Q34198 Sw:P51349 Sw:P37282 Sw:P235861		
Sw:P35862 Sw:Q05972 Tr:P93370		
/pseudo		
/codon_start=1		
/evidence=not experimental		
complement(24669. .25200)		
/note="match: GSS: Em:AQ504827"		
complement(25663. .26219)		
/note="match: GSS: Em:AQ309682"		
26085. .26326		
/note="match: GSS: Em:AQ463218"		
26086. .26569		
/note="match: GSS: Em:AQ347313"		
26106. .26574		
/note="match: GSS: Em:AQ405763"		
26294. .26464		
/note="L2 repeat: matches 2362. .2529 of consensus"		
27969. .28555		
/note="L2 repeat: matches 1685. .2333 of consensus"		
28558. .28868		
/note="MERS8B repeat: matches 23. .337 of consensus"		
29639. .29709		
/note="LTR33 repeat: matches 44. .119 of consensus"		
29715. .30122		
/note="LTR33 repeat: matches 102. .517 of consensus"		
30214. .30320		
/note="MER20 repeat: matches 2. .102 of consensus"		
30357. .30478		
/note="MER20 repeat: matches 98. .218 of consensus"		
30592. .30802		
/note="MIR repeat: matches 1. .216 of consensus"		
31220. .31393		
/note="MIR repeat: matches 56. .226 of consensus"		
32435. .32652		
/note="MIR repeat: matches 2. .261 of consensus"		
complement(join(32897. .33009,104107. .104314))		
/gene="dJ1007G16.2"		
complement(<32897. .>.33009)		
/gene="dJ1007G16.2"		
/note="exon borders derived from GENESCAN predictions; true borders could be further down		

```

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAATA 17
    |||||
Db 82046 TGTGGAAGGAATAATA 82062
    |||||

RESULT 9
AC079112/c 155888 bp DNA linear PRI 07-NOV-2001
LOCUS Homo sapiens BAC clone RP11-67K11 from 2, complete sequence.
DEFINITION AC079112
ACCESSION AC079112.4 GI:14091934
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 155888)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
MEDLINE
9847074
PUBMED
2 (bases 1 to 155888)
Goyea,E., Mishra,S. and Kozlowski,A.
The sequence of Homo sapiens BAC clone RP11-67K11
Unpublished
3 (bases 1 to 155888)
Waterston,R.H.
Direct Submission
Waterston,R.H.
Submitted (18-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 155888)
Waterston,R.H.
Direct Submission
Submitted (16-MAY-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 155888)
Waterston,R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 16, 2001 this sequence version replaced gi:13752170.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0067K11
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this

sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frerking,E., Tateo,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-30P5; the clone sequenced to the right is RP11-474G23. Actual start of this clone is at base position 1 of RP11-67K11; actual end is at base position 155888 of RP11-67K11.

#### FEATURES

Location/Qualifiers

```

1..155888
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-67K11"
/clone_lib="RP11-11"
328..526
/rpt_family="MER1_type"
538..645
/rpt_family="MIR"
1745..2080
/rpt_family="ERV1"
2082..2215
/rpt_family="Alu"
2219..2389
/rpt_family="ERV1"
2390..2821
/rpt_family="ERV1"
3746..3978
/rpt_family="MIR"
4181..4281
/rpt_family="MIR"
4536..4849
/rpt_family="Alu"
5603..5990
/rpt_family="MaLR"
5993..7584
/rpt_family="MaLR"
7585..7708
/rpt_family="Alu"
7709..7824
/rpt_family="MaLR"
7825..7905
/rpt_family="MaLR"
7907..8614
/rpt_family="L1"
8615..10087
/rpt_family="L1"
10088..10108
/rpt_family="AT-rich"
10172..10432
/rpt_family="Alu"
10433..10741
/rpt_family="MaLR"
11491..11646
/rpt_family="MIR"
12515..12833
/rpt_family="Alu"
12910..13181
/rpt_family="Alu"
13182..13483
/rpt_family="Alu"
13749..14156

```

```

/rept_family="L2"
14158. .14654
/rept_family="MaLR"
14763. .14928
/rept_family="MaLR"
14990. .15229
/rept_family="MaLR"
15940. .17923
/rept_family="L1"
18057. .18200
/rept_family="L1"
18806. .18917
/rept_family="CA-rich"
19295. .19331
/rept_family="AT-rich"
20065. .20084
/rept_family="(TC)n"
20091. .20135
/rept_family="(TG)n"
20136. .20179
/rept_family="CT-rich"
21523. .21686
/rept_family="MER1_type"
21829. .22040
/rept_family="MIR"
22370. .22664
/rept_family="L1"
22757. .22858
/rept_family="L1"
23117. .23428
/rept_family="Alu"
23565. .24609
/rept_family="L1"
25098. .25143
/rept_family="AT-rich"
25479. .25657
/rept_family="L2"
25658. .25929
/rept_family="Alu"
25930. .26003
/rept_family="L2"
26003. .26168
/rept_family="MIR"
26390. .26739
/rept_family="MER2_type"
26789. .27175
/rept_family="MaLR"
27258. .27303
/rept_family="(CA)n"
27536. .27604
/rept_family="MIR"
27589. .27627
/rept_family="L2"
28276. .28369
/rept_family="MIR"
28432. .28836
/rept_family="L2"
29950. .30197
/rept_family="MIR"
30283. .30327
/rept_family="(CA)n"
30417. .30565
/rept_family="ERVU"
31120. .31275

Query Match 100.0%; Score 17; DB 9; Length 155888;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGAGGAGTAATA 17
Db 126218 TTGGAGGAGTAATA 126202

```

AC021453 156165 bp DNA linear HTG 01-APR-2000  
Homo sapiens clone RP11-125C16, WORKING DRAFT SEQUENCE, 14  
unordered pieces.

AC021453  
AC021453.3 GI:7382318  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
1 (bases 1 to 156165)  
Homo sapiens, clone RP11-125C16  
Unpublished  
2 (bases 1 to 156165)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,  
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Landers,T., Lehoczy,J., Levine,R., Lied,C., Liu,G., Locke,K.,  
Macdonald,P., Marquis,N., McEwan,P., McDurk,A., McKernan,K.,  
McPheeters,R., Meldrim,J., Menes,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Olivat,T.M., Peterson,K.,  
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,W.

Direct Submission  
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 1, 2000 this sequence version replaced gi:6721267.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L5169  
Center clone name: 125 C.16  
----- Summary Statistics  
Sequencing vector: M13, M7815, 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 145142 bases at least Q40  
Consensus quality: 151009 bases at least Q30  
Consensus quality: 153345 bases at least Q20  
Insert size: 160000; agarose-fp  
Insert size: 154865; sum-of-contigs  
Quality coverage: 4.1 in Q20 bases; agarose-fp  
Quality coverage: 4.2 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1836: contig of 1836 bp in length  
\* 1837 1936: gap of 100 bp  
\* 1937 5796: contig of 3860 bp in length

```

* 5797 5896: gap of 100 bp
* 5897 7977: contig of 3901 bp in length
* 7978 9897: gap of 100 bp
* 9898 13919: contig of 4022 bp in length
* 13920 14019: gap of 100 bp
* 14020 17400: contig of 3381 bp in length
* 17401 17500: gap of 100 bp
* 17501 21253: contig of 3753 bp in length
* 21254 21353: gap of 100 bp
* 21354 31095: contig of 9743 bp in length
* 31097 31195: gap of 100 bp
* 31197 41465: contig of 10269 bp in length
* 41466 41565: gap of 100 bp
* 41566 50901: contig of 9336 bp in length
* 50902 51001: gap of 100 bp
* 51002 62041: contig of 11040 bp in length
* 62042 62141: gap of 100 bp
* 62142 75247: contig of 13106 bp in length
* 75248 75347: gap of 100 bp
* 75348 98974: contig of 14527 bp in length
* 98975 116362: contig of 100 bp
* 116363 116462: gap of 100 bp
* 116463 156165: contig of 39703 bp in length.

FEATURES             Location/Qualifiers
     source            1..156165
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clone="RP11-125C16"
                        /clone_lib="RPC1-11 Human Male BAC"
     misc_feature      1..1836
                        /note="assembly_fragment"
     misc_feature      1937..5796
                        /note="assembly_fragment"
     misc_feature      5897..9797
                        /note="assembly_fragment"
     misc_feature      9898..13919
                        /note="assembly_fragment"
     misc_feature      14020..17400
                        /note="assembly_fragment"
     misc_feature      17501..21253
                        /note="assembly_fragment"
                        clone_end:SP6
                        vector_side:right
     misc_feature      21354..31095
                        /note="assembly_fragment"
     misc_feature      31197..41465
                        /note="assembly_fragment"
     misc_feature      41566..50901
                        /note="assembly_fragment"
     misc_feature      51002..62041
                        /note="assembly_fragment"
     misc_feature      62142..75247
                        /note="assembly_fragment"
     misc_feature      75348..89974
                        /note="assembly_fragment"
                        clone_end:T7
                        vector_side:right
     misc_feature      89975..116362
                        /note="assembly_fragment"
     misc_feature      116463..156165
                        /note="assembly_fragment"

```

## ORIGIN

```

Query Match          100.0%; Score 17; DB 2; Length 156165;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Cy 1 TGTGGAGGAATAAATA 17
    |||||
Db 2341 TGTGGAGGAATAAATA 2357

```

## RESULT 11

```

AC100852      157980 bp      DNA      linear      PRI 29-AUG-2002
LOCUS      Homo sapiens chromosome 17, clone RP11-125C16, complete sequence.
DEFINITION      AC100852
ACCESSION      AC100852
VERSION      AC100852.2 GI:22539166
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 157980)
AUTHORS      Birren,B., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 17, clone RP11-125C16
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 157980)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kanat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamazares,R., Landers,R., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,K., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 157980)
Birn,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepe,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kellis,C., Landers,R., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 29, 2002 this sequence version replaced gi17048222.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

----- Project Information  
 Center project name: L21987  
 Center clone name: 125\_C\_16  
 -----

# FEATURES

Location/Qualifiers  
 1..157980  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-125C16"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 62..173  
 /rpt\_family="MIR"  
 complement(636..832)  
 /rpt\_family="MIR"  
 833..888  
 /rpt\_family="MER3"  
 complement(891..1207)  
 /rpt\_family="AluJb"  
 1208..1283  
 /rpt\_family="MER3"  
 1232..1236  
 /note="< 30 qual SNGL region"  
 complement(1303..2370)  
 /rpt\_family="L1PA4"  
 1330..1354  
 /note="<30 qual SNGL region"  
 2371..3213  
 /rpt\_family="L1PA4"  
 complement(3266..3408)  
 /rpt\_family="Charlie7"  
 3506..3538  
 /rpt\_family="L2"  
 3539..3653  
 /rpt\_family="MER47A"  
 complement(3654..3926)  
 /rpt\_family="AluSc"  
 3927..4141  
 /rpt\_family="MER47A"  
 4142..4438  
 /rpt\_family="AluY"  
 4439..4473  
 /rpt\_family="MER47A"  
 4474..4561  
 /rpt\_family="L2"  
 complement(4601..4700)  
 /rpt\_family="L2"  
 complement(4708..4845)  
 /rpt\_family="MERSA"  
 complement(4944..5104)  
 /rpt\_family="MIR"  
 complement(5112..5532)  
 /rpt\_family="L2"  
 complement(5546..5822)  
 /rpt\_family="AluJb"  
 complement(5823..5999)  
 /rpt\_family="L2"  
 6000..6216  
 /rpt\_family="AluJb"  
 complement(6217..6292)  
 /rpt\_family="L2"  
 6447..6752  
 /rpt\_family="AluJo"  
 7412..7472  
 /rpt\_family="A-rich"  
 7789..7907  
 /rpt\_family="(CGG)n"  
 8642..8683  
 /rpt\_family="T-rich"  
 8934..8978  
 /rpt\_family="(CCA)n"

repeat\_region 9084..9123  
 /rpt\_family="(CAG)n"  
 repeat\_region 10995..11017  
 /rpt\_family="AT-rich"  
 repeat\_region 11225..11326  
 /rpt\_family="L2"  
 repeat\_region complement(11372..11504)  
 /rpt\_family="MLT1B"  
 repeat\_region 13034..13209  
 /rpt\_family="MIR"  
 repeat\_region complement(15480..15590)  
 /rpt\_family="L2"  
 repeat\_region 15703..15988  
 /rpt\_family="AluSq"  
 repeat\_region complement(16005..16318)  
 /rpt\_family="L2"  
 repeat\_region 16412..16555  
 /rpt\_family="MER106B"  
 repeat\_region 16825..17099  
 /rpt\_family="AluJb"  
 repeat\_region 17241..17261  
 /rpt\_family="AT-rich"  
 repeat\_region 17280..17310  
 /rpt\_family="(TTTG)n"  
 repeat\_region complement(17366..17658)  
 /rpt\_family="AluSc"  
 repeat\_region 17707..17808  
 /rpt\_family="MIR"  
 repeat\_region 18250..18277  
 /rpt\_family="AT-rich"  
 repeat\_region complement(18576..18864)  
 /rpt\_family="MER33"  
 repeat\_region complement(18926..19077)  
 /rpt\_family="AluJo"  
 repeat\_region 20786..20952  
 /rpt\_family="AluJo"  
 repeat\_region 20962..21087  
 /rpt\_family="MIR"  
 repeat\_region complement(21088..21390)  
 /rpt\_family="AluSc"  
 repeat\_region complement(21419..21646)  
 /rpt\_family="AluJo"  
 repeat\_region 21647..21715  
 /rpt\_family="MIR"  
 repeat\_region complement(21969..22223)  
 /rpt\_family="L1ME4A"  
 repeat\_region complement(22288..22411)  
 /rpt\_family="MIR3"  
 repeat\_region complement(22950..23256)  
 /rpt\_family="AluSp"  
 repeat\_region 24607..24789  
 /rpt\_family="FAM"  
 repeat\_region 24846..25144  
 /rpt\_family="AluJb"  
 repeat\_region 25229..25254  
 /rpt\_family="AT-rich"  
 repeat\_region 25700..26104

Query Match 100.0%; Score 17; DB 9; Length 157980;  
 Best Local Similarity 100.0%; Pred.No.2.6e+02;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGAATAATA 17  
 |||||  
 Db 4546 TGTGGAAGGAATAATA 4562

RESULT 12  
 AC024563  
 LOCUS AC024563 158541 bp DNA linear PRI 14-JUL-2002  
 DEFINITION Homo sapiens chromosome 19 clone CTC-451A6, complete sequence.  
 ACCESSION AC024563  
 VERSION AC024563.6 GI:21747460

```

KEYWORDS      HTG.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 158541)
JOURNAL       DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE      Unpublished
AUTHORS       2 (bases 1 to 158541)
TITLE         DOE Joint Genome Institute.
JOURNAL       Direct Submission
REFERENCE      Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
AUTHORS       Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE         3 (bases 1 to 158541)
JOURNAL       DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE      Direct Submission
AUTHORS       DOE Joint Genome Institute
TITLE         Submitted (01-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
JOURNAL       Drive, Walnut Creek, CA 94598, USA
REFERENCE      5 (bases 1 to 158541)
AUTHORS       DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE         Direct Submission
JOURNAL       Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS       Drive, Walnut Creek, CA 94598, USA
TITLE         On Jul 14, 2002 this sequence version replaced gi:15887308.
JOURNAL       Draft Sequence Produced by DOE Joint Genome Institute
COMMENT       www.jgi.doe.gov
              Finishing completed at Stanford Human Genome Center
              Quality: Phrap Quality >=40 99.9% of Sequence;
              Estimated Total Number of Errors is 0.7.
FEATURES      Location/Qualifiers
               source
               1..158541
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="19"
               /clone="CFC-451A6"

ORIGIN
Query Match      100.0%; Score 17; DB 9; Length 158541;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGTGGAAGGATAATA 17
Db      9645 TGTGGAAGGATAATA 9661

RESULT 13
AC009108/c
LOCUS      AC009108      168656 bp      DNA      linear      PRI 29-OCT-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-46309, complete sequence.
ACCESSION  AC009108
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE   1 (bases 1 to 168656)
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      DOE Joint Genome Institute, Stanford Human Genome Center and Los
JOURNAL    Alamos National Laboratory.
REFERENCE   Unpublished
AUTHORS    2 (bases 1 to 168656)
TITLE      DOE Joint Genome Institute.
JOURNAL    Direct Submission
AUTHORS    DOE Joint Genome Institute.
TITLE      Direct Submission

KEYWORDS      HTG.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 168656)
JOURNAL       DOE Joint Genome Institute, Stanford Human Genome Center and Los
REFERENCE      Alamos National Laboratory.
AUTHORS       Direct Submission
TITLE         Unpublished
JOURNAL       Submitted (29-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell
AUTHORS       Drive, Walnut Creek, CA 94598, USA
TITLE         On Oct 29, 2002 this sequence version replaced gi:13786306.
JOURNAL       Draft Sequence Produced by DOE Joint Genome Institute
COMMENT       www.jgi.doe.gov
              Finishing completed at Stanford Human Genome Center and Los Alamos
              National Laboratory
              www.sngc.stanford.edu
              Quality: Phrap Quality >=40 99.9% of Sequence;
              Estimated Total Number of Errors is 0.2.
FEATURES      Location/Qualifiers
               source
               1..168656
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="16"
               /clone="RP11-46309"

ORIGIN
Query Match      100.0%; Score 17; DB 9; Length 168656;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGTGGAAGGATAATA 17
Db      78374 TGTGGAAGGATAATA 78358

RESULT 14
AC107725/c
LOCUS      AC107725      169226 bp      DNA      linear      HTG 21-AUG-2002
DEFINITION Mus musculus clone RP23-477015, WORKING DRAFT SEQUENCE, 5 ordered
              pieces.
ACCESSION  AC107725
VERSION    AC107725.4 GI:2381821
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE   1 (bases 1 to 169226)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL    Birren,B., Nusbaum,C. and Lander,E.
REFERENCE   2 (bases 1 to 169226)
AUTHORS    Unpublished
TITLE      Mus musculus, clone RP23-477015
JOURNAL    Unpublished
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
              Brown,A., Camarata,J., Campopiano,A., Chang,J.J., Chazaro,B.,
              Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
              Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,
              Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
              Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
              Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
              Landers,T., Lenocky,J., Levine,R., Liu,G., Maclean,C.,
              Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
              McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
              Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
              Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
              Peterson,K., Phukhang,P., Pierre,N., Pollara,V., Raymond,C.,
              Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
              Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
              Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
              Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
              Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
              Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169226)
Birnre, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hegos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:20163106.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19818
Center clone name: 477 O.15
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 168312 bases at least Q40
Consensus quality: 168665 bases at least Q30
Consensus quality: 168771 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 168826; sum-of-contigs
Quality coverage: 12.4 in Q20 bases; agarose-fp
Quality coverage: 12.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequences as soon as it is available and
* the accession number will be preserved.
1 39640: contig of 39640 bp in length
39740: gap of 100 bp
39741 56981: contig of 17241 bp in length
56982 57081: gap of 100 bp
57082 74209: contig of 17128 bp in length
74210 74309: gap of 100 bp
74310 141150: contig of 66841 bp in length
141151 141250: gap of 100 bp
141251 169226: contig of 27976 bp in length.
Location/Qualifiers
1 .169226
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-477O15"
FEATURES
source
/clone_lib="RPCI-23 Female Mouse BAC"
1.39640
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
39741..56981
/note="assembly_fragment"
57082..74209
/note="assembly_fragment"
74310..141150
/note="assembly_fragment"
141251..169226
/note="assembly_fragment"
clone_end:T7
vector_side:right"
ORIGIN
Query Match 100.0%; Score 17; DB 2; Length 169226;
Best Local Similarity 100.0%; Pred.No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGGAAGGATAAATA 17
|||||
Db 79687 TGTGGAAGGATAAATA 79671
|||||
RESULT 15
AC025272/c 170749 bp DNA linear PRI 03-FEB-2001
LOCUS Homo sapiens chromosome 15 clone RP11-548M13 map 15q21.3, complete
sequence.
AC025272
AC025272.6 GI:12658003
VERSION
HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 170749)
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A., Madan, A.,
Nesbitt, R., Traicoff, R. and Hood, L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
2 (bases 1 to 170749)
Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B.,
Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G.,
James, R., Kaur, A., Madan, A., Owen, M.P., Ratcliffe, A., Shaffer, T.
and Hood, L.
Direct Submission
Submitted (08-MAR-2000) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
3 (bases 1 to 170749)
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A., Madan, A.,
Nesbitt, R., Traicoff, R. and Hood, L.
Direct Submission
Submitted (03-FEB-2001) Multimegabase Sequencing Center, Institute
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
98105, USA
On Feb 3, 2001 this sequence version replaced gi:8272670.
----- Genome Center
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Assembly program: Phrap; version 0.990399
Note: Data from overlapping clones AC068723 [Drafting center:

```

UWWSCL, AC073100 [Drafting center: WUGSC] and AC011912 [Drafting center: UWWSCL] were added for finishing

FEATURES

source

Location/Qualifiers

1. .170749  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21.3"  
/clone="RP11-548M13"  
/clone\_lib="RPCI human BAC library 11"  
/note="This clone overlaps RP11-126E3 AC068723, RP11-643P18 AC073100 and RP11-547D13 AC011912. Data from overlapping BACs were added and the consensus sequence determined from RP11-548M13 to the extent possible."  
misc\_feature  
1. .137140  
/notes="overlap with RP11-643P18, AC073100"  
misc\_feature  
1. .24070  
/notes="overlap with RP11-126E3, AC068723"  
38050. .38070  
/notes="low quality data"  
49230. .49320  
/notes="low quality data"  
54520. .54540  
/notes="low quality data"  
80700. .80900  
/notes="low quality data"  
81160. .81210  
/notes="low quality data"  
misc\_feature  
116818. .170749  
/notes="overlap with RP11-547D13, AC011912"  
121430. .121440  
/notes="low quality data"

ORIGIN

Query Match 100.0%; Score 17; DB 9; Length 170749;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTGAAGGAATAATA 17  
|||||  
Db 41977 TTGTGAAGGAATAATA 41961

Search completed: March 25, 2004, 12:49:49  
Job time : 119.83 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 07:40:23 ; Search time 17.642 Seconds

(without alignments)  
4093.601 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375

Perfect score: 17

Sequence: 1 tgggaaggaataaata 17

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	17	100.0	6458	6	ABK48984 Genomic D
2	16	94.1	405	4	AAK73688 Human imm
3	16	94.1	406	4	AAK73689 Human imm
4	16	94.1	418	4	AAK58790 Human imm
5	16	94.1	1807	2	AAK74041 Soybean t
6	16	94.1	9124	4	AAK73690 Human imm
7	16	94.1	21619	4	AAK73691 Human imm
8	16	94.1	21619	4	AAK73691 Human rep
9	16	94.1	21619	4	AAK73691 Human ova
10	15.4	90.6	50	6	ABK48984 Genomic D
11	15.4	90.6	215	2	AAK73688 Human imm
12	15.4	90.6	601	9	AAK73689 Human imm
13	15.4	90.6	601	9	AAK58790 Human imm
14	15.4	90.6	602	9	AAK74041 Soybean t
15	15.4	90.6	657	4	AAK73690 Human imm
16	15.4	90.6	943	5	AAK73691 Human imm
17	15.4	90.6	1140	3	AAK73691 Human rep
18	15.4	90.6	1555	4	AAK73691 Human ova
19	15.4	90.6	1780	6	AAK73691 Human imm
20	15.4	90.6	2336	6	AAK73691 Human imm
21	15.4	90.6	2339	4	AAK73691 Human rep
22	15.4	90.6	2755	4	AAK73691 Human ova
23	15.4	90.6	3022	6	AAK73691 Human imm

24	15.4	90.6	4584	4	ABL05142 Drosophil
25	15.4	90.6	6021	6	ABK48984 Genomic D
26	15.4	90.6	6052	6	AAK73688 Human imm
27	15.4	90.6	8197	6	ABL34515 Human met
28	15.4	90.6	8197	6	ABL70542 Human met
29	15.4	90.6	16281	4	ABK42480 Genomic s
30	15.4	90.6	16281	4	AAK70314 Human imm
31	15.4	90.6	16281	8	ABK42480 Genomic s
32	15.4	90.6	16285	4	ABK42481 Genomic s
33	15.4	90.6	16285	4	ABK42482 Genomic s
34	15.4	90.6	16285	4	AAK70315 Human imm
35	15.4	90.6	16285	4	AAK70316 Human imm
36	15.4	90.6	16285	8	ABK42480 Genomic s
37	15.4	90.6	16285	8	ABK42481 Genomic s
38	15.4	90.6	17294	6	ABK42482 Genomic s
39	15.4	90.6	26000	7	AAK73688 Human imm
40	15.4	90.6	50000	3	AAK73689 Human imm
41	15.4	90.6	52116	4	AAK73689 Human imm
42	15.4	90.6	52116	6	AAK73689 Human imm
43	15.4	90.6	55008	4	AAK80495 Human imm
44	15.4	90.6	164702	7	ACF62730 Cancer ba
45	15.4	90.6	164702	7	ADB20845 MRP1 base

## ALIGNMENTS

## RESULT 1

ABK48984

ID ABK48984 standard; DNA; 6458 BP.

XX AC ABK48984;

XX DT 02-JUL-2002 (first entry)

XX DE Genomic DNA encoding human transcription factor FOXC2.

XX KW Transcription factor; FOXC2, antidiabetic; anorectic; antilipaeamic;

XX KW cardiovascular; FOXC2 polypeptide expression modulator; FOXC2 promoter;

XX KW diabetes mellitus; obesity; hypercholesterolaemia; dyslipidaemia;

XX KW cardiovascular disease; gene; ds.

XX OS Homo sapiens.

XX FH Key

XX exon

FT 1..215

FT /tag= b

FT /note= "First exon according to the alternative

FT transcript"

FT 1..186

FT /tag= a

FT /tag= 187..215

FT /tag= c

FT /note= "Region coding for 5' part of alternative protein"

FT 215..216

FT /tag= d

FT /note= "Alternative first exon splice site"

FT misc\_difference 216..475

FT /tag= e

FT /note= "Fragment of the FOXC2 enhancer. Specifically

FT claimed in claim 16"

FT misc\_difference 223..231

FT /tag= f

FT /note= "Fragment of the FOXC2 enhancer. Specifically

FT claimed in claim 12"

FT misc\_difference 359..375

FT /tag= g

FT /note= "Fragment of the FOXC2 enhancer. Specifically

FT claimed in claim 13"

FT misc\_difference 378..402

FT /tag= h

FT /note= "Fragment of the FOXC2 enhancer. Specifically

FT claimed in claim 14"

FT misc\_difference 403..423  
FT /\*tag= i  
FT /note= "Fragment of the FOXC2 enhancer. Specifically  
FT claimed in claim 15"  
FT misc\_difference 1250..2235  
FT /\*tag= k  
FT /note= "Fragment of the FOXC2 promoter. Specifically  
FT claimed in claim 3"  
FT misc\_difference 1250..1749  
FT /\*tag= j  
FT /note= "Fragment of the FOXC2 promoter. Specifically  
FT claimed in claim 2"  
FT misc\_difference 1692..1703  
FT /\*tag= l  
FT /note= "Fragment of the FOXC2 promoter. Specifically  
FT claimed in claim 1"  
FT exon 1746..4629  
FT /\*tag= h  
FT /note= "First exon according to the published form of the  
FT transcript"  
FT 5'UTR 1746..2234  
FT /\*tag= m  
FT CDS 2235..3740  
FT /\*tag= o  
FT /product= "FOXC2"  
FT /note= "Transcription factor"  
FT 2448..2735  
FT /\*tag= p  
FT /note= "Region coding for DNA-binding domain"  
FT 2516..4629  
FT /\*tag= r  
FT /note= "Second exon according to the alternative  
FT transcript"  
FT misc\_feature 2516..3740  
FT /\*tag= q  
FT /note= "Portion of polypeptide used in alternative  
FT transcript"  
FT 3'UTR 3741..4629  
FT /\*tag= s  
XX WO200227008-A1.  
XX  
XX  
PD 04-APR-2002.  
XX  
XX 26-SEP-2001; 2001WO-SE002098.  
XX  
XX 26-SEP-2000; 2000SE-00003435.  
XX 10-OCT-2000; 2000US-0238897P.  
XX 09-NOV-2000; 2000SE-00004102.  
XX  
XX (BIOV-) BIOVITRUM AB.  
XX  
XX Enerbaeck S, Krook K, Rondahl L, Wasserman WW;  
XX WPI; 2002-352129/38.  
XX P-ESDB; AAU79816.  
XX  
XX An isolated FOXC2 promoter region that modulates the expression of a  
XX FOXC2 polypeptide is useful for treating obesity and type II diabetes  
XX mellitus.  
XX  
XX Claim 3; Page 33-38; 62pp; English.  
XX  
XX The invention describes an isolated human FOXC2 promoter region. The  
XX mammalian FOXC2 promoter region for the modulation of FOXC2 polypeptide  
XX expression has therapeutic value in treating type II diabetes mellitus,  
XX obesity, hypercholesterolemia, other cardiovascular diseases or  
XX dyslipidaemias. This sequence encodes the human transcription factor  
XX FOXC2, the promoter of which is described in the invention  
XX  
XX Sequence 6458 BP; 1443 A; 1826 C; 1715 G; 1474 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 17; DB 6; Length 6458;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGTGGAGGAATAAATA 17  
DB 359 TGTGGAGGAATAAATA 375  
RESULT 2  
AAK73688/c  
ID AAK73688 standard; DNA; 405 BP.  
XX  
AC AAK73688;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28500.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0225759P.  
PR 22-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-022824P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.

```
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237032P.
PR 02-OCT-2000; 2000US-0237033P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240950P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250161P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251857P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 28500; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 405 BP; 126 A; 73 C; 50 G; 156 T; 0 U; 0 Other;
XX
XX Query Match 94.1%; Score 16; DB 4; Length 405;
XX Best Local Similarity 100.0%; Pred. No. 3.4e+02;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TGTGGAAGGATAAAT 16
XX |||||
XX 361 TGTGGAAGGATAAAT 346
XX
XX RESULT 3
XX AAK73689/C
XX ID AAK73689 standard; DNA; 406 BP.
XX
XX AAK73689;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28501.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX
```

OS Homo sapiens.  
XX WO200157182-A2.  
XX 09-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001354.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 08-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232337P.  
PR 14-SEP-2000; 2000US-0232338P.  
PR 14-SEP-2000; 2000US-0232339P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235494P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-024121P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246537P.  
PR 08-NOV-2000; 2000US-0246538P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR (HUMA-) HUMAN GENOME SCI INC.  
XX  
PR

```

XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 28501; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX SQ Sequence 406 BP; 126 A; 74 C; 50 G; 156 T; 0 U; 0 Other;

Query Match 94.1%; Score 16; DB 4; Length 406;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAGGGAATAAT 16
Db 362 TGTGGAGGGAATAAT 347

RESULT 4
AAK58790/c
ID AAK58790 standard; cDNA; 418 BP.
XX AC AAK58790;
XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:3850.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW Cytostatic; gene therapy; vaccine; metastasis; ss.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001354.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184684P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209457P.
XX PR 28-JUN-2000; 2000US-0214866P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225277P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

```

PR 20-OCT-2000; 2000US-0241B26P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246529P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249247P.  
 PR 17-NOV-2000; 2000US-0249248P.  
 PR 17-NOV-2000; 2000US-0249249P.  
 PR 17-NOV-2000; 2000US-0249250P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0256147P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX P-PSDB; AAK86009.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.

XX Claim 1; SEQ ID NO 3850; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the

CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 418 BP; 124 A; 77 C; 55 G; 157 T; 0 U; 5 Other;

Query Match 94.1%; Score 16; DB 4; Length 418;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGAGAGGATAAAT 16

DB 371 TGTGAGAGGATAAAT 356

RESULT 5

AAAT74041  
 ID AAT74041 standard; cDNA to mRNA; 1807 BP.

XX AC AAT74041;

XX 17-SEP-1997 (first entry)

XX DE Soybean thiol protease D3-alpha encoding cDNA.

XX KW Soybean; cotyledon; germinate; transformant; Escherichia coli;  
 XX recombinant expression vector; ds.

XX OS Glycine max.

XX FH Key Location/Qualifiers

XX CDS 68..1459

XX FT /\*tag= a

XX FT /\*product= "Thiol\_protease"

XX FT sig\_peptide 68..1460

XX FT /\*tag= b

XX FT mat\_peptide 461..1459

XX FT /\*tag= c

XX FT /\*product= "D3-alpha"

XX PN JP09121870-A.

XX XX

XX PD 13-MAY-1997.

XX XX

XX PF 28-DEC-1995; 95JP-00353931.

XX XX

XX PR 29-DEC-1994; 94JP-00340399.

XX PR 30-AUG-1995; 95JP-00245279.

XX XX

XX PA (AJIN ) AJINOMOTO KK.

XX WPI; 1997-314230/29.

XX P-PSDB; AAK19541.

XX PT New thiol protease and corresponding DNA sequence - used in the  
 PT preparation of the thiol protease.

XX PS Claim 1; Page 10-12; 17pp; Japanese.

XX CC The present sequence encodes a novel thiol protease D3-alpha derived from  
 CC germinated soybean cotyledon. A method for the production of thiol  
 CC protease using Escherichia coli transformed with recombinant expression  
 CC vectors containing a DNA sequence encoding thiol protease has been  
 CC produced. The method can be used to prepare large amounts of thiol  
 CC protease with high efficiency

XX SQ Sequence 1807 BP; 500 A; 387 C; 449 G; 471 T; 0 U; 0 Other;

Query Match

94.1%; Score 16; DB 2; Length 1807;

```
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GTGAAGGAATAAATA 17
Db 566 GTGAAGGAATAAATA 581

RESULT 6
AAK73690/c
ID AAK73690 standard; DNA; 9124 BP.
AC AAK73690;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28502.
DE
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001354.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR
XX 04-FEB-2000; 2000US-0180628P.
PR
XX 24-FEB-2000; 2000US-0184664P.
PR
XX 02-MAR-2000; 2000US-0186350P.
PR
XX 16-MAR-2000; 2000US-0189874P.
PR
XX 17-MAR-2000; 2000US-0190076P.
PR
XX 18-APR-2000; 2000US-0196123P.
PR
XX 19-MAY-2000; 2000US-0205515P.
PR
XX 07-JUN-2000; 2000US-0209467P.
PR
XX 28-JUN-2000; 2000US-0214886P.
PR
XX 30-JUN-2000; 2000US-0215135P.
PR
XX 07-JUL-2000; 2000US-0216647P.
PR
XX 07-JUL-2000; 2000US-0216880P.
PR
XX 11-JUL-2000; 2000US-0217487P.
PR
XX 11-JUL-2000; 2000US-0217496P.
PR
XX 14-JUL-2000; 2000US-0218290P.
PR
XX 26-JUL-2000; 2000US-0220963P.
PR
XX 26-JUL-2000; 2000US-0220964P.
PR
XX 14-AUG-2000; 2000US-0224518P.
PR
XX 14-AUG-2000; 2000US-0224519P.
PR
XX 14-AUG-2000; 2000US-0225213P.
PR
XX 14-AUG-2000; 2000US-0225214P.
PR
XX 14-AUG-2000; 2000US-0225266P.
PR
XX 14-AUG-2000; 2000US-0225267P.
PR
XX 14-AUG-2000; 2000US-0225268P.
PR
XX 14-AUG-2000; 2000US-0225270P.
PR
XX 14-AUG-2000; 2000US-0225447P.
PR
XX 14-AUG-2000; 2000US-0225757P.
PR
XX 14-AUG-2000; 2000US-0225758P.
PR
XX 14-AUG-2000; 2000US-0225759P.
PR
XX 18-AUG-2000; 2000US-0226279P.
PR
XX 22-AUG-2000; 2000US-0226681P.
PR
XX 22-AUG-2000; 2000US-0226686P.
PR
XX 22-AUG-2000; 2000US-0227182P.
PR
XX 23-AUG-2000; 2000US-0227009P.
PR
XX 30-AUG-2000; 2000US-0228924P.
PR
XX 01-SEP-2000; 2000US-0229287P.
PR
XX 01-SEP-2000; 2000US-0229343P.
PR
XX 01-SEP-2000; 2000US-0229344P.
PR
XX 01-SEP-2000; 2000US-0229345P.
PR
XX 05-SEP-2000; 2000US-0229509P.
PR
XX 05-SEP-2000; 2000US-0229513P.
PR
XX 06-SEP-2000; 2000US-0230437P.
PR
XX 06-SEP-2000; 2000US-0230438P.
PR
XX 08-SEP-2000; 2000US-0231242P.
PR
XX 08-SEP-2000; 2000US-0231243P.
PR
XX 08-SEP-2000; 2000US-0231244P.
PR
XX 08-SEP-2000; 2000US-0231413P.
PR
XX 08-SEP-2000; 2000US-0231414P.
PR
XX 08-SEP-2000; 2000US-0232080P.
PR
XX 08-SEP-2000; 2000US-0232081P.
PR
XX 12-SEP-2000; 2000US-0231968P.
PR
XX 14-SEP-2000; 2000US-0232397P.
PR
XX 14-SEP-2000; 2000US-0232398P.
PR
XX 14-SEP-2000; 2000US-0232399P.
PR
XX 14-SEP-2000; 2000US-0232400P.
PR
XX 14-SEP-2000; 2000US-0232401P.
PR
XX 14-SEP-2000; 2000US-0233063P.
PR
XX 14-SEP-2000; 2000US-0233064P.
PR
XX 14-SEP-2000; 2000US-0233065P.
PR
XX 21-SEP-2000; 2000US-0234223P.
PR
XX 21-SEP-2000; 2000US-0234274P.
PR
XX 25-SEP-2000; 2000US-0234997P.
PR
XX 25-SEP-2000; 2000US-0234998P.
PR
XX 26-SEP-2000; 2000US-0235484P.
PR
XX 27-SEP-2000; 2000US-0235834P.
PR
XX 27-SEP-2000; 2000US-0235836P.
PR
XX 29-SEP-2000; 2000US-0236327P.
PR
XX 29-SEP-2000; 2000US-0236367P.
PR
XX 29-SEP-2000; 2000US-0236368P.
PR
XX 29-SEP-2000; 2000US-0236369P.
PR
XX 29-SEP-2000; 2000US-0236370P.
PR
XX 02-OCT-2000; 2000US-0236802P.
PR
XX 02-OCT-2000; 2000US-0237037P.
PR
XX 02-OCT-2000; 2000US-0237038P.
PR
XX 02-OCT-2000; 2000US-0237039P.
PR
XX 13-OCT-2000; 2000US-0237040P.
PR
XX 13-OCT-2000; 2000US-0239335P.
PR
XX 13-OCT-2000; 2000US-0239337P.
PR
XX 20-OCT-2000; 2000US-0240960P.
PR
XX 20-OCT-2000; 2000US-0241221P.
PR
XX 20-OCT-2000; 2000US-0241785P.
PR
XX 20-OCT-2000; 2000US-0241786P.
PR
XX 20-OCT-2000; 2000US-0241787P.
PR
XX 20-OCT-2000; 2000US-0241808P.
PR
XX 20-OCT-2000; 2000US-0241809P.
PR
XX 01-NOV-2000; 2000US-0241828P.
PR
XX 08-NOV-2000; 2000US-0244617P.
PR
XX 08-NOV-2000; 2000US-0246474P.
PR
XX 08-NOV-2000; 2000US-0246475P.
PR
XX 08-NOV-2000; 2000US-0246476P.
PR
XX 08-NOV-2000; 2000US-0246477P.
PR
XX 08-NOV-2000; 2000US-0246478P.
PR
XX 08-NOV-2000; 2000US-0246523P.
PR
XX 08-NOV-2000; 2000US-0246524P.
PR
XX 08-NOV-2000; 2000US-0246525P.
PR
XX 08-NOV-2000; 2000US-0246526P.
PR
XX 08-NOV-2000; 2000US-0246527P.
PR
XX 08-NOV-2000; 2000US-0246528P.
PR
XX 08-NOV-2000; 2000US-0246532P.
PR
XX 08-NOV-2000; 2000US-0246609P.
PR
XX 08-NOV-2000; 2000US-0246610P.
PR
XX 08-NOV-2000; 2000US-0246611P.
PR
XX 17-NOV-2000; 2000US-0249213P.
PR
XX 17-NOV-2000; 2000US-0249207P.
PR
XX 17-NOV-2000; 2000US-0249208P.
PR
XX 17-NOV-2000; 2000US-0249209P.
PR
XX 17-NOV-2000; 2000US-0249210P.
PR
XX 17-NOV-2000; 2000US-0249211P.
PR
XX 17-NOV-2000; 2000US-0249212P.
PR
XX 17-NOV-2000; 2000US-0249213P.
PR
XX 17-NOV-2000; 2000US-0249214P.
PR
XX 17-NOV-2000; 2000US-0249215P.
PR
XX 17-NOV-2000; 2000US-0249216P.
PR
XX 17-NOV-2000; 2000US-0249217P.
PR
XX 17-NOV-2000; 2000US-0249218P.
PR
XX 17-NOV-2000; 2000US-0249244P.
```

PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249298P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251038P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
FA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Disclosure; SEQ ID NO 28502; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54950 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX  
XX Query Match 94.1%; Score 16; DB 4; Length 9124;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGTGGAGGATAAAT 16  
Db 5310 TGTGGAGGATAAAT 5295  
RESULT 7  
AAK73691/C  
ID AAK73691 standard; DNA; 21619 BP.  
XX  
AC AAK73691;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:28503.  
XX  
XX Human; immune; hematopoietic; immune/hematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX

OS Homo sapiens.  
XX WO200157182-A2.  
XX  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198113P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.



PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-02559678P.  
(HUNA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX Disclosure; SEQ ID NO 28503; 3071pp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I/  
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
CC represent sequences used in the exemplification of the present invention  
XX SQ Sequence 21619 BP; 7128 A; 3520 C; 3719 G; 7252 T; 0 U; 0 Other;  
Query Match 94.1%; Score 16; DB 4; Length 21619;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGTGGAAGGAATAAT 16  
Db 5310 TGTGGAAGGAATAAT 5295  
RESULT 8  
AAL07315/c  
ID AAL07315 standard; DNA; 21619 BP.  
XX AC AAL07315;  
XX DT 21-NOV-2001 (first entry)  
XX DE Human reproductive system related antigen DNA SEQ ID NO: 10003.  
XX KW Human; reproductive system related antigen; reproductive system disorder;  
XX OS cancer; gene therapy; ds.  
XX OS Homo sapiens.  
XX PN WO200155320-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001339.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218390P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226691P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240600P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-02559678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.

Disclosure; SEQ ID NO 10003; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention

Sequence 21619 BP; 7128 A; 3520 C; 3719 G; 7252 T; 0 U; 0 Other;

Query Match 94.1%; Score 16; DB 4; Length 21619;

Best Local Similarity 100.0%; Pred. No. 4.4e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGCAATAAAT 16  
D5 5310 TGTGGAAGCAATAAAT 5295  
RESULT 9  
ABA08182/c  
ID ABA08182 standard; DNA; 21619 BP.  
XX  
AC ABA08182;  
XX  
DT 11-JAN-2002 (first entry)  
XX Human ovarian and breast cancer associated polynucleotide SEQ ID NO 977.  
XX  
DE  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO20015325-A2.  
XX  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001345.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190075P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225266P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 14-AUG-2000; 2000US-0225759P.  
XX 18-AUG-2000; 2000US-0226279P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226686P.  
XX 23-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0228287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229503P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240360P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246809P.  
PR 08-NOV-2000; 2000US-0246810P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249247P.  
 PR 17-NOV-2000; 2000US-0249248P.  
 PR 17-NOV-2000; 2000US-0249249P.  
 PR 17-NOV-2000; 2000US-0249250P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250161P.  
 PR 05-DEC-2000; 2000US-0250331P.  
 PR 05-DEC-2000; 2000US-0250332P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0251989P.  
 PR 06-DEC-2000; 2000US-0251719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251898P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-488786/53.

XX New isolated ovarian and/or breast cancer related nucleic acids and  
 PT polypeptides, useful for diagnosing, treating and/or preventing human  
 PT diseases and disorders, particularly ovarian and/or breast cancer.

PS Disclosure; SEQ ID NO 977; 577bp + Sequence Listing; English.

XX The invention relates to novel genes (ABA07454-ABA08224) and proteins  
 CC (ABB10743-ABB10980) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 21619 BP; 7128 A; 3520 C; 3719 G; 7252 T; 0 U; 0 Other;

Query Match 94.1%; Score 16; DB 4; Length 21619;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAGGATATAAT 16

DB 5310 TGTGGAGGATATAAT 5295

RESULT 10

ABZ01147  
 ID ABZ01147 standard; DNA; 50 BP.

XX ABZ01147;

AC ABZ01147;

XX 09-JAN-2003 (first entry)

XX Human leukocyte gene expression profiling probe SEQ ID NO 1138.

DE T7; leukocyte; gene expression profiling; allograft rejection;

XX

KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;  
 KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;  
 SS.

XX Homo sapiens.

XX WO200257414-A2.

XX 25-JUL-2002.

XX 22-OCT-2001; 2001WO-US047856.

XX 20-OCT-2000; 2000US-0241994P.

PR 08-JUN-2001; 2001US-0296764P.

XX (BIOC-) BIOCARDIA INC.

XX Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;  
 PI Ly N, Woodward R, Quattermost T, Johnson F;

XX WPI; 2002-636525/68.

XX New system for leukocyte expression profiling, diagnosing a disease, or  
 PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis  
 PT or congestive heart failure, comprises diagnostic oligonucleotides.

XX Claim 1; Page 361; Opp; English.

XX The invention relates to a system for detecting gene expression, which  
 CC comprises one or two isolated DNA molecules that detect expression of a  
 CC gene, where the gene corresponds to any of 8143 oligonucleotides  
 CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful  
 CC for leukocyte expression profiling. It is particularly useful for  
 CC diagnosing a disease, monitoring (rate of) progression of a disease,  
 CC predicting therapeutic outcome, determining prognosis for a patient,  
 CC predicting disease complications in an individual or monitoring response  
 CC to treatment in an individual. The diseases include cardiac allograft  
 CC rejection, kidney allograft rejection, liver allograft rejection,  
 CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,  
 CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection  
 XX Sequence 50 BP; 13 A; 5 C; 13 G; 19 T; 0 U; 0 Other;

Query Match 90.6%; Score 15.4; DB 6; Length 50;

Best Local Similarity 94.1%; Pred. No. 5.8e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGGAGGATATAATA 17

DB 27 TGTGGAGGATATAATA 43

RESULT 11

AAT20169  
 ID AAT20169 standard; cDNA to mRNA; 215 BP.

XX AAT20169;

XX 23-JUL-1996 (first entry)

XX Human gene signature HUMGS01315.

XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX WO9514772-A1.

XX 01-JUN-1995.

XX 11-NOV-1994; 94WO-JP001916.

XX PR 12-NOV-1993; 93JP-00355504.  
XX PA (MATSU) MATSUBARA K.  
XX PA (OKUBO) OKUBO K.  
XX PI Matsubara K, Okubo K;  
XX DR WPI; 1995-206931/27.  
XX PT Single-stranded DNA for identifying gene signatures - isolated from 3'-  
XX PT directed human cDNA library that reflects relative abundance of corresp.  
XX PT mRNA in specific human tissues.  
XX PS Claim 1; Page 577; 2245pp; Japanese.  
XX CC A single-stranded DNA (or its complementary strand or the corresp. double  
XX CC -stranded DNA) which comprises one of the 7837 "GS" sequences given in  
XX CC AR19001-T26837 and which is able to hybridize to part of human genomic  
XX CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were  
XX CC obtained from 3'-directed cDNA libraries prepared from various human  
XX CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using  
XX CC poly(T) as the sole primer. Since the 3'- untranslated sequence is unique  
XX CC to a particular mRNA species, almost all the 3'-oriented cDNAs hybridize  
XX CC with specific mRNAs. Each library is constructed so as to reflect  
XX CC accurately the relative abundance of different mRNAs in the particular  
XX CC tissue from which it was derived. The appearance frequency of a given GS  
XX CC in a cDNA library can be determined (esp. using primers and probes  
XX CC derived from the GS sequences) as a means of diagnosing abnormal cell  
XX CC function or for recognising different cell types  
XX SQ Sequence 215 BP; 71 A; 32 C; 36 G; 72 T; 0 U; 4 Other;  
Query Match 90.6%; Score 15.4; DB 2; Length 215;  
Best Local Similarity 94.1%; Pred. No. 6.4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGTGGAGGAGTAATAAATA 17  
178 TGTGGAGGAGTAATAAATA 194  
Db  
RESULT 12  
ADD17249  
ID ADD17249 standard; DNA; 601 BP.  
AC ADD17249;  
XX DT 15-JAN-2004 (first entry)  
XX DE DNA (SeqID 1317) that confers an altered visual phenotype in plants.  
XX KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;  
XX KW bleaching; etching; wet leaf; stunting; elongation; texture;  
XX KW agronomic trait; growth regulation; dwarf variety; insect resistance;  
XX KW heat stress; transgenic.  
XX OS Unidentified.  
XX OS WO2003020741-A1.  
XX PN 13-MAR-2003.  
XX PD 30-AUG-2002; 2002WO-US027880.  
XX PR 31-AUG-2001; 2001US-0316326P.  
XX PA (DOWC ) DOW CHEM CO.  
XX PA (DOWC ) DOW AGROSCIENCES LLC.  
XX PI Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;  
XX PS WPI; 2003-300858/29.  
XX DR

XX PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae  
XX PT sativa, Saccharomyces cerevisiae, and Papaver rhoeas, useful for  
XX PT conferring altered visual phenotypes in plants.  
XX PS Claim 1; SEQ ID NO 1317; 517pp; English.  
XX CC This invention relates to the identification and isolation of novel  
XX CC nucleic acid molecules that confer altered visual phenotypes in plants.  
XX CC Specifically, it refers to modifications of plant architecture and/or  
XX CC leaf surface features in plants, such as chlorotic, bleaching, etching,  
XX CC wet leaf, stunting, elongation and texture phenotypes, which are thought  
XX CC will be agronomic traits beneficial to the farmer. As such, these novel  
XX CC phenotypes can affect growth regulation i.e. useful for creating dwarf  
XX CC varieties, exhibit resistance to insects or heat stress, confer changes  
XX CC in pigment content such that plants have enhanced vitamin production or  
XX CC delayed senescence and also for example produce plants that control the  
XX CC production of ethylene. Furthermore, the present invention comprises  
XX CC generating transgenic plants, as well as reproducibly altering the visual  
XX CC phenotype of plant seeds, plant tissues and plant cells containing the  
XX CC polynucleotides described herein. This polynucleotide is a homologue of a  
XX CC DNA sequence that confers an altered visual phenotype when expressed in  
XX CC plants, the method of the invention.  
XX SQ Sequence 601 BP; 178 A; 105 C; 148 G; 170 T; 0 U; 0 Other;  
Query Match 90.6%; Score 15.4; DB 9; Length 601;  
Best Local Similarity 94.1%; Pred. No. 6.9e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGTGGAGGAGTAATAAATA 17  
425 TGTGGAGGAGTAATAAATA 441  
Db  
RESULT 13  
ADD17846  
ID ADD17846 standard; DNA; 601 BP.  
AC ADD17846;  
XX DT 15-JAN-2004 (first entry)  
XX DE DNA (SeqID 1914) that confers an altered visual phenotype in plants.  
XX KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;  
XX KW bleaching; etching; wet leaf; stunting; elongation; texture;  
XX KW agronomic trait; growth regulation; dwarf variety; insect resistance;  
XX KW heat stress; transgenic.  
XX OS Unidentified.  
XX OS WO2003020741-A1.  
XX PN 13-MAR-2003.  
XX PD 30-AUG-2002; 2002WO-US027880.  
XX PR 31-AUG-2001; 2001US-0316326P.  
XX PA (DOWC ) DOW CHEM CO.  
XX PA (DOWC ) DOW AGROSCIENCES LLC.  
XX PI Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;  
XX PS WPI; 2003-300858/29.  
XX DR  
XX PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae  
XX PT sativa, Saccharomyces cerevisiae, and Papaver rhoeas, useful for  
XX PT conferring altered visual phenotypes in plants.  
XX PS Claim 1; SEQ ID NO 1914; 517pp; English.  
XX XX

varieties, exhibit resistance to insects or heat stress, confer changes in pigment content such that plants have enhanced vitamin production or delayed senescence and also for example produce plants that control the production of ethylene. Furthermore, the present invention comprises generating transgenic plants, as well as reproducibly altering the visual phenotype of plant seeds, plant tissues and plant cells containing the polynucleotides described herein. This polynucleotide is a homologue of a DNA sequence that confers an altered visual phenotype when expressed in plants, the method of the invention.

Sequence	602 BP; 182 A; 106 C; 145 G; 169 T; 0 U; 0 Other;
Query Match	90.6%;
Score	15.4; DB 9;
Length	60
Best Local Similarity	94.1%;
Pred. No.	6.9e+02;

Qy 1 TGTGGAAGGAATAAATA 17  
Db 426 TGTGGAAGGAATAAAGA 442

RESULT 15	
AAH31558	
ID	AAH31558 standard; cDNA; 657 BP.
XX	
XX	AAH31558;
XX	
XX	
DT	30-JUL-2001 (first entry)
XX	
XX	
DE	Human olfactory receptor polynucleotide, SEQ ID NO: 131.
XX	
XX	Human; olfactory receptor; OR; primary scent determination;
KW	secondary scent determination; polypeptide library; odour receptor;
KW	scor profile; scent fingerprint; scent representation; ss.

30-JUL-2001 (first entry)

Human olfactory receptor polynucleotide, SEQ ID NO: 131.

Human; olfactory receptor; OR; primary scent determination;  
 secondary scent determination; polypeptide library; odour receptor;  
 scent profile. scent fingerprint; scent representation: ss.

100

xx DNA (SeqID 419) that confers an altered visual phenotype in plants.  
DE  
xx  
xx ds; visual phenotype; plant; architecture; leaf surface; chlorotic;  
KW bleaching; etching; wet leaf; stunting; elongation; texture;  
KW agronomic trait; growth regulation; dwarf variety; insect resistant  
KW heat stress; transgenic.

08-OCT-1999; 99US-0158615P.  
24-FEB-2000; 2000US-0184809P.  
(DIGI-) DIGISCENTS.  
(YEDA ) YEDA RES & DEV CO LTD.  
Bellenson J, Smith D, Lancet  
WPI; 2001-290713/30.

AA New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists.  
PT  
PT

PS Claim 1: Fig 2: 1857pp; English.

The present sequence is one of a number of isolated polynucleotides which encode polypeptides involved in olfactory sensation. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents and the identification of combinations of odour receptors that are involved in detecting such secondary scents. This enables the construction of a scent representation (also called a scent fingerprint or scent profile), which may be used to re-create and edit scents. Libraries of olfactory receptors are useful for determining the interaction pattern of a composition with the receptors, and can be used for determining differences in the olfactory faculties of different individuals.

Query Match 90.6%; Score 15.4; DB 4; Length 657;  
Best Local Similarity 94.1%; Fred. No. 6.9e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGGAGGAATAAATA 17  
|||  
Db 102 TCGGAGGAATAAATA 118

Search completed: March 25, 2004, 10:25:16  
Job time : 21.642 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:53:09 ; Search time 128.47 seconds

(without alignments)  
3951.570 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375

Perfect score: 17

Sequence: 1 tgtggaaggaataaata 17

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gsa\_hum.\*

18: em\_gsa\_inv.\*

19: em\_gsa\_pln.\*

20: em\_gsa\_vrt.\*

21: em\_gsa\_fun.\*

22: em\_gsa\_mam.\*

23: em\_gsa\_mus.\*

24: em\_gsa\_pro.\*

25: em\_gsa\_rtd.\*

26: em\_gsa\_phg.\*

27: em\_gsa\_vrl.\*

28: gb\_gsa1.\*

29: gb\_gsa2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17	100.0	362	9	AV654470
C 2	17	100.0	411	12	BJ577286
C 3	17	100.0	445	28	AQ080438
4	17	100.0	510	28	AQ247792

5	17	100.0	552	28	AQ055707
C 6	17	100.0	684	12	BJ572614
C 7	17	100.0	773	12	BJ571448
C 8	17	100.0	781	12	BJ570143
C 9	17	100.0	791	12	BJ567603
C 10	17	100.0	814	12	BJ570146
C 11	17	100.0	815	12	BJ569616
C 12	17	100.0	831	12	BJ578691
C 13	16.2	95.3	1179	9	AL526913
14	16	94.1	217	10	BB330439
15	16	94.1	224	10	BB331990
16	16	94.1	258	10	BB335604
17	16	94.1	266	28	AZ327704
C 18	16	94.1	287	10	BB334643
C 19	16	94.1	310	10	BB394077
C 20	16	94.1	340	9	AI850188
21	16	94.1	358	10	BE654556
C 22	16	94.1	400	10	BE953799
C 23	16	94.1	407	28	AQ216184
24	16	94.1	418	9	AU254064
C 25	16	94.1	431	28	AQ562634
26	16	94.1	437	28	AQ265092
27	16	94.1	456	29	CE748440
C 28	16	94.1	477	14	CB239719
C 29	16	94.1	488	12	BG838806
30	16	94.1	493	12	BM094320
31	16	94.1	502	10	BE655174
32	16	94.1	519	12	BG651500
33	16	94.1	520	13	BU884047
34	16	94.1	542	10	BE347121
35	16	94.1	543	10	BF425191
36	16	94.1	549	10	AW831547
37	16	94.1	553	13	BQ626094
38	16	94.1	563	13	BU898223
C 39	16	94.1	566	13	BU800227
40	16	94.1	568	12	BM085956
41	16	94.1	572	12	BI469425
42	16	94.1	573	12	BI786379
C 43	16	94.1	586	12	BG838815
44	16	94.1	598	10	BF597251
45	16	94.1	601	10	AW203570

#### ALIGNMENTS

RESULT 1  
AV654470/c  
LOCUS AV654470 GLC Homo sapiens cDNA clone GLCDB09 3', mRNA linear EST 15-JAN-2002  
DEFINITION AV654470 GLC Homo sapiens cDNA clone GLCDB09 3', mRNA sequence.  
ACCESSION AV654470  
VERSION AV654470.1 GI:9875484  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 362)  
XU,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,  
Shao.H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,  
Xiao.K., Lu.G., Fu.G., Zhong.M., Xu,S., Gu,W., Huang,W., Zhao,X.,  
Hu.G., Gu,J., Chen,Z. and Han,Z.  
Insight into hepatocellular carcinogenesis at transcriptome level  
by comparing gene expression profiles of hepatocellular carcinoma  
with those of corresponding noncancerous liver  
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

Contact: Zequang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)



```

Fax: 86-21-50801922
Email: hantg@cgsc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..362
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="GLCWB03"
      /tissue_type="corresponding non cancerous liver tissue"
      /dev_stage="Adult"
      /lab_host="SOLR"
      /clone_lib="GLC"
      /notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
      XhoI"

ORIGIN
  Query Match      100.0%; Score 17; DB 9; Length 362;
  Best Local Similarity 100.0%; Pred. No. 2.3e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17
Db 227 TGTGGAAGGAATAAATA 211

RESULT 2
BJ577286/c
LOCUS
DEFINITION
  BJ577286 Ipomoea nil mixture of flower and flower bud Ipomoea nil
  CNA clone jm34k01 3, mRNA sequence.
VERSION
  BJ577286
KEYWORDS
  BJ577286.1 GI:27259114
SOURCE
  Ipomoea nil (Japanese morning glory)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamiales; Solanales; Convolvulaceae; Ipomoea.
REFERENCE
  1 (bases 1 to 411)
  Nitsuka, E., Shinozaki, K., Hayashizaki, Y., Kohara, Y. and Iida, S.
  Hoshino, A., Seki, M., Shin-i, T., Carninci, P., Kamiya, A., Shiraki, T.,
  ESTs of Japanese morning glory
  Unpublished (2002)
  Contact: Tadasi Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@gene.nig.ac.jp.
FEATURES
  source
    1..411
      /organism="Ipomoea nil"
      /mol_type="mRNA"
      /cultivar="Tokyo-kokei standard"
      /db_xref="taxon:35893"
      /clone="jm34k01"
      /tissue_type="mixture of flower and flower bud"
      /clone_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN
  Query Match      100.0%; Score 17; DB 12; Length 411;
  Best Local Similarity 100.0%; Pred. No. 2.2e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17
Db 73 TGTGGAAGGAATAAATA 57

RESULT 3
.AQ080438

```

```

LOCUS
DEFINITION
  AQ080438 445 bp DNA linear GSS 20-AUG-1998
  CIT-HSP-2382B13.TF CIT-HSP Homo sapiens genomic clone 2382B13,
  genomic survey sequence.
ACCESSION
  AQ080438
VERSION
  AQ080438.1 GI:3441622
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 445)
  Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
  Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
  Venter, J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Other GSSs: CIT-HSP-2382B13.TR
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends
FEATURES
  source
    1..445
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="2382B13"
      /sex="Male"
      /cell_type="Sperm"
      /clone_lib="CIT-HSP"
      /notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
      HindIII"
ORIGIN
  Query Match      100.0%; Score 17; DB 28; Length 445;
  Best Local Similarity 100.0%; Pred. No. 2.2e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17
Db 411 TGTGGAAGGAATAAATA 427

RESULT 4
AQ247792
LOCUS
DEFINITION
  HS_2063_A2_G01_MR CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=2063 Col=2 Row=M, Genomic survey
  sequence.
ACCESSION
  AQ247792
VERSION
  AQ247792.1 GI:3697974
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 510)
  Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., and
  Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
  Hood, L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  99380589

```

10449764  
 PUBMED  
 COMMENT  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Sequence Tagged Connector  
 Plate: 2063 row: M column: 2  
 Class: BAC ends  
 High quality sequence stop: 510.  
 Location/Qualifiers  
 1. 510  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=2063 Col=2 Row=M"  
 /sex="male"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /notes="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

ORIGIN  
 Query Match 100.0%; Score 17; DB 28; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGAATAAATA 17  
 |||||  
 Db 182 TGTGGAAGGAATAAATA 198

RESULT 5  
 AQ055707  
 LOCUS  
 DEFINITION  
 CIT-HSP-2342D16, TR CIT-HSP Homo sapiens genomic clone 2342D16,  
 genomic survey sequence.  
 ACCESSION  
 AQ055707  
 VERSION  
 AQ055707.1 GI:3352313  
 KEYWORDS  
 GSS.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 652)  
 AUTHORS  
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
 Simon,M. and Venter,J.C.  
 TITLE  
 Use of a random BAC End Sequence Database for Sequence-Ready Map  
 Building (1998)  
 JOURNAL  
 Unpublished (1998)  
 COMMENT  
 Other GSSs: CIT-HSP-2342D16.TF  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdamadms@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13 Reverse  
 Class: BAC ends.  
 Location/Qualifiers  
 1. 652  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2342D16"  
 /sex="Male"  
 /cell\_type="Sperm"

/clone\_lib="CIT-HSP"  
 /note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2:  
 HindIII"

ORIGIN  
 Query Match 100.0%; Score 17; DB 28; Length 652;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGAATAAATA 17  
 |||||  
 Db 200 TGTGGAAGGAATAAATA 216

RESULT 6  
 BJ572614/c  
 LOCUS  
 DEFINITION  
 BJ572614 Ipomoea nil mixture of flower and flower bud Ipomoea nil  
 CDNA clone jm21e10 3', mRNA sequence.  
 ACCESSION  
 BJ572614  
 VERSION  
 BJ572614.1 GI:27254442  
 KEYWORDS  
 EST.  
 SOURCE  
 Ipomoea nil (Japanese morning glory)  
 ORGANISM  
 Ipomoea nil  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Convolvulaceae; Ipomoea.  
 1 (bases 1 to 684)  
 AUTHORS  
 Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,  
 Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.  
 TITLE  
 ESTs of Japanese morning glory  
 JOURNAL  
 Unpublished (2002)  
 COMMENT  
 Contact: Tadabu Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshini@genes.nig.ac.jp.  
 Location/Qualifiers  
 1. 684  
 /organism="Ipomoea nil"  
 /mol\_type="mRNA"  
 /cultivar="Tokyo-kokei standard"  
 /db\_xref="taxon:35883"  
 /clone="jm21e10"  
 /tissue\_type="mixture of flower and flower bud"  
 /clone\_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN  
 Query Match 100.0%; Score 17; DB 12; Length 684;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGGAAGGAATAAATA 17  
 |||||  
 Db 130 TGTGGAAGGAATAAATA 114

RESULT 7  
 BJ571448/c  
 LOCUS  
 DEFINITION  
 BJ571448 Ipomoea nil mixture of flower and flower bud Ipomoea nil  
 CDNA clone jm17k24 3', mRNA sequence.  
 ACCESSION  
 BJ571448  
 VERSION  
 BJ571448.1 GI:27253276  
 KEYWORDS  
 EST.  
 SOURCE  
 Ipomoea nil (Japanese morning glory)  
 ORGANISM  
 Ipomoea nil  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Convolvulaceae; Ipomoea.  
 1 (bases 1 to 773)

AUTHORS Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,  
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.  
ESTs of Japanese morning glory  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
Location/Qualifiers  
source  
1..773  
/organism="Ipomoea nil"  
/mol\_type="mRNA"  
/cultivar="Tokyo-kokei standard"  
/db\_xref="taxon:35883"  
/c1one="jm17k24"  
/tissue\_type="mixture of flower and flower bud"  
/clone\_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN  
Query Match 100.0%; Score 17; DB 12; Length 773;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17  
|||||

Db 171 TGTGGAAGGAATAAATA 155  
|||||

RESULT 8  
BJ570143/c  
LOCUS  
DEFINITION BJ570143 Ipomoea nil mixture of flower and flower bud Ipomoea nil  
CDNA clone jm7102 3', mRNA sequence.  
ACCESSION BJ570143  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Ipomoea nil (Japanese morning glory)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Convolvulaceae; Ipomoea.

REFERENCE  
1 (bases 1 to 781)  
Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,  
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.  
ESTs of Japanese morning glory  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
Location/Qualifiers  
source  
1..781  
/organism="Ipomoea nil"  
/mol\_type="mRNA"  
/cultivar="Tokyo-kokei standard"  
/db\_xref="taxon:35883"  
/c1one="jm7102"  
/tissue\_type="mixture of flower and flower bud"  
/clone\_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN  
Query Match 100.0%; Score 17; DB 12; Length 781;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17  
|||||

Db 72 TGTGGAAGGAATAAATA 56  
|||||

RESULT 9  
BJ567603/c  
LOCUS  
DEFINITION BJ567603 Ipomoea nil mixture of flower and flower bud Ipomoea nil  
CDNA clone jm14m18 3', mRNA sequence.  
ACCESSION BJ567603  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Ipomoea nil (Japanese morning glory)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Convolvulaceae; Ipomoea.

REFERENCE  
1 (bases 1 to 791)  
Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,  
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.  
ESTs of Japanese morning glory  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
Location/Qualifiers  
source  
1..791  
/organism="Ipomoea nil"  
/mol\_type="mRNA"  
/cultivar="Tokyo-kokei standard"  
/db\_xref="taxon:35883"  
/c1one="jm14m18"  
/tissue\_type="mixture of flower and flower bud"  
/clone\_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN  
Query Match 100.0%; Score 17; DB 12; Length 791;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGAATAAATA 17  
|||||

Db 144 TGTGGAAGGAATAAATA 128  
|||||

RESULT 10  
BJ570146/c  
LOCUS  
DEFINITION BJ570146 Ipomoea nil mixture of flower and flower bud Ipomoea nil  
CDNA clone jm7105 3', mRNA sequence.  
ACCESSION BJ570146  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Ipomoea nil (Japanese morning glory)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Convolvulaceae; Ipomoea.

REFERENCE  
1 (bases 1 to 814)  
Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,  
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.  
ESTs of Japanese morning glory  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

```

FEATURES
  source
    Location/Qualifiers
      1..814
        /organism="Ipomoea nil"
        /mol_type="mRNA"
        /cultivar="Tokyo-kokei standard"
        /db_xref="taxon:35883"
        /clone="jm7105"
        /tissue_type="mixture of flower and flower bud"
        /clone_lib="Ipomoea nil mixture of flower and flower bud"
ORIGIN
  Query Match      100.0%; Score 17; DB 12; Length 814;
  Best Local Similarity 100.0%; Pred. No. 2.1e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAGGATAAATA 17
    |||
Db 98 TGTGGAGGATAAATA 82
    |||

RESULT 11
BJ569616/c
LOCUS
DEFINITION
  BJ569616 Ipomoea nil mixture of flower and flower bud Ipomoea nil
  CDS clone jm6c03 3', mRNA sequence.
ACCESSION
  BJ569616
VERSION
  BJ569616.1 GI:27251444
KEYWORDS
  EST.
SOURCE
  Ipomoea nil (Japanese morning glory)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    asterids; lamiales; Solanales; Convolvulaceae; Ipomoea.
  1 (bases 1 to 815)
  Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
  Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
  ESTs of Japanese morning glory
  Unpublished (2002)
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
  Location/Qualifiers
    1..831
      /organism="Ipomoea nil"
      /mol_type="mRNA"
      /cultivar="Tokyo-kokei standard"
      /db_xref="taxon:35883"
      /clone="jm39b09"
      /tissue_type="mixture of flower and flower bud"
      /clone_lib="Ipomoea nil mixture of flower and flower bud"
ORIGIN
  Query Match      100.0%; Score 17; DB 12; Length 831;
  Best Local Similarity 100.0%; Pred. No. 2.1e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAGGATAAATA 17
    |||
Db 71 TGTGGAGGATAAATA 55
    |||

RESULT 13
AL526913
LOCUS
DEFINITION
  AL526913 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
  CDNA clone CSODC017YM10 5-PRIME, mRNA sequence.
ACCESSION
  AL526913
VERSION
  AL526913.2 GI:31064771
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  1 (bases 1 to 1179)
  Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  Full-length cDNA libraries and normalization
  Unpublished (2001)
  On Feb 13, 2001 this sequence version replaced gi:12790406.
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  Library was constructed by Life Technologies, a division of
  Invitrogen. Contact : Feng Liang Email: fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
  Faraday Avenue Genoscope sequence ID : CSODC017YG05QPI.
  Location/Qualifiers
    1..1179
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CSODC017YM10"
      /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo (dT)
      primer. Five prime end enriched, double-strand cDNA was
      digested with Not I and cloned into the Not I and EcoR V
FEATURES
  source
    Location/Qualifiers
      1..814
        /organism="Ipomoea nil"
        /mol_type="mRNA"
        /cultivar="Tokyo-kokei standard"
        /db_xref="taxon:35883"
        /clone="jm6c03"
        /tissue_type="mixture of flower and flower bud"
        /clone_lib="Ipomoea nil mixture of flower and flower bud"
ORIGIN
  Query Match      100.0%; Score 17; DB 12; Length 815;
  Best Local Similarity 100.0%; Pred. No. 2.1e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGGAGGATAAATA 17
    |||
Db 115 TGTGGAGGATAAATA 99
    |||

RESULT 12
BJ578691/c
LOCUS
DEFINITION
  BJ578691 Ipomoea nil mixture of flower and flower bud Ipomoea nil
  CDS clone jm39b09 3', mRNA sequence.
ACCESSION
  BJ578691
VERSION
  BJ578691.1 GI:27260519
KEYWORDS
  EST.

```

sites of the PCWSPORT 6 vector. Library was normalized."

# ORIGIN

Query Match. 95.3%; Score 16.2; DB 9; Length 1179;  
 Best Local Similarity 88.2%; Pred. No. 4.6e+03;  
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGTGGAAGGATAAATA 17  
 |||||:|||||:  
 Db 405 TGTGGAAGGARTAAATW 421

RESULT 14  
 BB330439  
 LOCUS BB330439 RIKEN full-length enriched, 6 days neonate medulla  
 DEFINITION oblongata Mus musculus cDNA clone B730013P22 3', mRNA sequence.  
 BB330439  
 ACCESSION BB330439.1 GI:9039202  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 217)  
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,  
 Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,  
 Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,  
 Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,  
 Shibata,K., Shibata,Y., Shigemoto,Y., Suzuki,H., Tagawa,A.,  
 Takahashi,F., Tomimaga,N., Toyoda,T., Tsunoda,Y., Watanabe,S.,  
 Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,  
 Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and  
 Hayashizaki,Y.  
 RIKEN Mouse ESTs (Konno,H., et al.)  
 Unpublished (2000)

# TITLE

JOURNAL RIKEN Mouse ESTs (Konno,H., et al.)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,  
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermotabilization and thermoactivation of thermostable enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Kitsuunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
 Okazaki,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details

Location/Qualifiers  
 1. 217  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="B730013P22"  
 /tissue\_type="medulla oblongata"  
 /dev\_stage="6 days neonate"  
 /lab\_host="DH10B"

# FEATURES

## source

1. 217  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="B730013P22"  
 /tissue\_type="medulla oblongata"  
 /dev\_stage="6 days neonate"  
 /lab\_host="DH10B"

/clone\_lib="RIKEN full-length enriched, 6 days neonate  
 medulla oblongata"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 20.0 and subtraction to Rot = 459.0. Second  
 strand cDNA was prepared with the primer adaptor of  
 sequence [5' GAGAGAGATCTCGAGTAATTAATTAATCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified pluescript KS(+) after bulk excision from Lambda  
 PLC I."

# ORIGIN

Query Match 94.1%; Score 16; DB 10; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+03;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGTGGAAGGATAAATA 16  
 |||||:|||||:  
 Db 99 TGTGGAAGGATAAATA 114

# RESULT 15

BB331990  
 LOCUS BB331990 RIKEN full-length enriched, 6 days neonate medulla  
 DEFINITION oblongata Mus musculus cDNA clone B730041N08 3', mRNA sequence.  
 BB331990  
 ACCESSION BB331990.1 GI:9040753  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 224)  
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,  
 Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,  
 Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,  
 Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,  
 Shibata,K., Shibata,Y., Shigemoto,Y., Suzuki,H., Tagawa,A.,  
 Takahashi,F., Tomimaga,N., Toyoda,T., Tsunoda,Y., Watanabe,S.,  
 Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,  
 Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and  
 Hayashizaki,Y.  
 RIKEN Mouse ESTs (Konno,H., et al.)  
 Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,  
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermotabilization and thermoactivation of thermostable enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Kitsuunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
 Okazaki,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details

# TITLE

JOURNAL RIKEN Mouse ESTs (Konno,H., et al.)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,  
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermotabilization and thermoactivation of thermostable enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Kitsuunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
 Okazaki,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details

## source

1. 217  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="B730013P22"  
 /tissue\_type="medulla oblongata"  
 /dev\_stage="6 days neonate"  
 /lab\_host="DH10B"

1. 217  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="B730013P22"  
 /tissue\_type="medulla oblongata"  
 /dev\_stage="6 days neonate"  
 /lab\_host="DH10B"

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 13-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

source

Location/Qualifiers

```

1..224
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B730041N08"
/tissue_type="medulla oblongata"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 6 days neonate medulla oblongata"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGAGAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
```

## ORIGIN

```

Query Match          94.1%; Score 16; DB 10; Length 224;
Best Local Similarity 100.0%; Pred. No. 6.4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TGTGGAGGGAATAAT 16
        |||||
Db       106 TGTGGAGGGAATAAT 121
```

Search completed: March 25, 2004, 15:30:09  
Job time : 133.47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:55:14 ; Search time 3.34382 Seconds  
(without alignments)  
2821.370 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_359\_375  
Perfect score: 17  
Sequence: 1 tgggaaggaataaata 17

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/pctus COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.4	90.6	26000	4	US-09-843-376-10
2	15	88.2	675	4	US-09-456-090A-99
3	14.4	84.7	313	1	US-08-396-452-1
4	14.4	84.7	313	3	US-08-169-119-1
5	14.4	84.7	391	4	US-08-621-976-11175
6	14.4	84.7	1056	1	US-09-500-651-1
7	14.4	84.7	1056	1	US-08-813-591-1
8	14.4	84.7	1225	4	US-09-398-412B-1
9	14.4	84.7	38844	4	US-09-734-675-3
10	14.4	84.7	152331	3	US-09-128-155-16
11	14.4	84.7	176373	3	US-09-138-155-17
12	14	82.4	234	4	US-08-485-050A-121
13	14	82.4	1127	2	US-08-392-546C-4
14	14	82.4	16592	4	US-08-956-171E-53
15	14	82.4	36741	3	US-09-301-665-3
16	13.8	81.2	76	4	US-09-743-871B-26
17	13.8	81.2	84	4	US-09-568-816A-7
18	13.8	81.2	492	3	US-08-328-111-335
19	13.8	81.2	510	3	US-08-081-590A-4
20	13.8	81.2	593	3	US-09-328-111-533
21	13.8	81.2	814	3	US-08-998-416-353
22	13.8	81.2	945	4	US-09-107-532A-3633
23	13.8	81.2	1024	4	US-09-414-276-5
24	13.8	81.2	1044	4	US-09-716-129-24
25	13.8	81.2	1110	4	US-09-414-276-2
26	13.8	81.2	1182	4	US-09-107-532A-2750
27	13.8	81.2	1256	4	US-09-743-871B-9

C	28	13.8	81.2	1773	4	US-09-107-532A-1028	Sequence 1028, Ap
	29	13.8	81.2	2114	4	US-09-568-816A-1	Sequence 1, Appl
	30	13.8	81.2	2190	4	US-08-887-534A-29	Sequence 29, Appl
	31	13.8	81.2	2190	4	US-09-527-431-29	Sequence 29, Appl
C	32	13.8	81.2	2518	4	US-09-743-871B-10	Sequence 10, Appl
C	33	13.8	81.2	2739	2	US-08-951-871-1	Sequence 1, Appl
C	34	13.8	81.2	3252	4	US-08-672-749-9	Sequence 9, Appl
C	35	13.8	81.2	3668	2	US-08-792-055-2	Sequence 2, Appl
	36	13.8	81.2	3742	4	US-09-016-434-1370	Sequence 1370, Ap
	37	13.8	81.2	5173	1	US-08-242-677-1	Sequence 1, Appl
	38	13.8	81.2	5473	4	US-08-620-312D-260	Sequence 260, App
	39	13.8	81.2	6040	4	US-10-204-708-70	Sequence 70, Appl
C	40	13.8	81.2	6320	4	US-09-620-312D-135	Sequence 135, App
	41	13.8	81.2	6574	4	US-09-221-017B-1097	Sequence 1097, Ap
	42	13.8	81.2	8351	1	US-08-198-446B-14	Sequence 14, Appl
	43	13.8	81.2	8351	2	US-08-870-693-14	Sequence 14, Appl
C	44	13.8	81.2	9062	4	US-08-956-171E-85	Sequence 85, Appl
C	45	13.8	81.2	15363	4	US-08-961-527-139	Sequence 139, App

ALIGNMENTS

RESULT 1  
US-09-843-376-10  
Sequence 1, Appl  
Patent No. 6566132  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF INTERFERON GAMMA RECEPTOR 1 EXPRESSION  
FILE REFERENCE: KIS-0234  
CURRENT APPLICATION NUMBER: US/09/843.376  
CURRENT FILING DATE: 2001-04-26  
NUMBER OF SEQ ID NOS: 88  
SEQ ID NO 10  
LENGTH: 26000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-843-376-10

Query Match 90.6%; Score 15.4; DB 4; Length 26000;  
Best Local Similarity 94.1%; Pred. No. 95;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTGAAGCAATAAATA 17  
DB 13569 TGTGAAGCAATAAATA 13585

RESULT 2  
US-09-456-090A-99  
Sequence 99, Appl  
Patent No. 6680209  
GENERAL INFORMATION:  
APPLICANT: Buechler, Joe  
APPLICANT: Valkirs, Gunars  
APPLICANT: Gray, Jeff  
APPLICANT: Lomborg, Nils  
TITLE OF INVENTION: HUMAN ANTIBODIES AS DIAGNOSTIC REAGENTS  
FILE REFERENCE: 020015-000200US  
CURRENT APPLICATION NUMBER: US/09/456.090A  
CURRENT FILING DATE: 1999-12-06  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 99  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(675)

```

; OTHER INFORMATION: M2-20H
US-09-456-090A-99
;
; Query Match      88.2%; Score 15; DB 4; Length 675;
; Best Local Similarity 100.0%; Pred. No. 96;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      3 TGGGAAGGAATAAATA 17
;         |||||
;         162 TGGGAAGGAATAAATA 176
;
; Db
;
; RESULT 3
; US-08-396-452-1/c
; Sequence 1, Application US/08396452
; Patent No. 5820863
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: Method of Preparation and Use for Zona
; TITLE OF INVENTION: Pellucida Antigens and Antibodies for Sterilization and
; TITLE OF INVENTION: Contraception
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,452
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/055,831
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M.
; TELEPHONE: 713-850-0909
; TELEX: 713-850-0165
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..313
; OTHER INFORMATION: /standard_name= "Rabbit cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-396-452-1
;
; Query Match      84.7%; Score 14.4; DB 1; Length 313;
; Best Local Similarity 93.8%; Pred. No. 1.7e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 TGTGAAGGAATAAATA 16
;         |||||
;         98 TGTGAAGGAATAAATA 83
;
; Db
;
; RESULT 4
; US-09-169-119-1/c
; Sequence 1, Application US/09169119
; Patent No. 6264953
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: Method of Preparation and Use for Zona
; TITLE OF INVENTION: Pellucida Antigens and Antibodies for Sterilization and
; Contraception
;
; QY      1 TGTGAAGGAATAAATA 16
;         |||||
;         98 TGTGAAGGAATAAATA 83
;
; Db
;
; Query Match      84.7%; Score 14.4; DB 4; Length 391;
; Best Local Similarity 93.8%; Pred. No. 1.8e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 TGTGAAGGAATAAATA 16
;         |||||
;         160 TGTGAAGGAATAAATA 175
;
; Db
```

```

; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/169,119
; FILING DATE: 09-Oct-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M.
; TELEPHONE: 713-850-0909
; TELEX: 713-850-0165
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..313
; OTHER INFORMATION: /standard_name= "Rabbit cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-169-119-1
;
; Query Match      84.7%; Score 14.4; DB 3; Length 313;
; Best Local Similarity 93.8%; Pred. No. 1.7e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 TGTGAAGGAATAAATA 16
;         |||||
;         98 TGTGAAGGAATAAATA 83
;
; Db
;
; RESULT 5
; US-09-621-976-11175
; Sequence 11175, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 11175
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-621-976-11175
;
; Query Match      84.7%; Score 14.4; DB 4; Length 391;
; Best Local Similarity 93.8%; Pred. No. 1.8e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 TGTGAAGGAATAAATA 16
;         |||||
;         160 TGTGAAGGAATAAATA 175
;
; Db
```



## RESULT 6

US-09-500-651-1  
; Sequence 1, Application US/09500651  
; GENERAL INFORMATION:  
; APPLICANT: ASANO, MINAO  
; APPLICANT: KAWAI, MISAKO  
; APPLICANT: MIWA, TETSUYA  
; APPLICANT: NIO, NORIKI  
; TITLE OF INVENTION: AMINOPEPTIDASE GX, AND A METHOD OF  
; TITLE OF INVENTION: HYDROLYSING A PROTEIN WITH THE SAME  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/500.651  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/813,591  
; FILING DATE:  
; APPLICATION DATA:  
; APPLICATION NUMBER: JP 030458/1997  
; FILING DATE: 14-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 10-845-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1056 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1056  
; OTHER INFORMATION: /note= "METHOD OF DETERMINING THE  
; OTHER INFORMATION: CHARACTERISTICS: P"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..9  
; OTHER INFORMATION: /note= "INSERTION SEQUENCE, METHOD  
; OTHER INFORMATION: OF DETERMINING THE CHARACTERISTICS: E"  
US-09-500-651-1

Query Match 84.7%; Score 14.4; DB 1; Length 1056;  
Best Local Similarity 93.8%; Pred. No. 2e+02; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGGAAGGAATAAATA 17  
|||  
Db 436 GTAGAAGGAATAAATA 451

## RESULT 7

US-08-813-591-1  
; Sequence 1, Application US/08813591  
; Patent No. 5824534  
; GENERAL INFORMATION:  
; APPLICANT: ASANO, MINAO  
; APPLICANT: KAWAI, MISAKO  
; APPLICANT: MIWA, TETSUYA  
; APPLICANT: NIO, NORIKI  
; TITLE OF INVENTION: AMINOPEPTIDASE GX, AND A METHOD OF  
; TITLE OF INVENTION: HYDROLYSING A PROTEIN WITH THE SAME  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/813,591  
; FILING DATE: 07-MAR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 051848/1996  
; FILING DATE: 08-MAR-1996  
; APPLICATION DATA:  
; APPLICATION NUMBER: JP 030458/1997  
; FILING DATE: 14-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 10-845-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1056 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1056  
; OTHER INFORMATION: /note= "METHOD OF DETERMINING THE  
; OTHER INFORMATION: CHARACTERISTICS: P"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..9  
; OTHER INFORMATION: /note= "INSERTION SEQUENCE, METHOD  
; OTHER INFORMATION: OF DETERMINING THE CHARACTERISTICS: E"  
US-08-813-591-1

Query Match 84.7%; Score 14.4; DB 1; Length 1056;  
Best Local Similarity 93.8%; Pred. No. 2e+02; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGGAAGGAATAAATA 17  
|||  
Db 436 GTAGAAGGAATAAATA 451

## RESULT 8

US-09-398-412B-1/c  
; Sequence 1, Application US/09398412B  
; Patent No. 6680380  
; GENERAL INFORMATION:  
; APPLICANT: Timans, Jacqueline C.  
; TITLE OF INVENTION: Nucleic acids encoding mammalian interleukin-1zeta, related reagents and uses thereof  
; FILE REFERENCE: methods  
; CURRENT APPLICATION NUMBER: US/09/398,412B  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: US 60/100948  
; PRIOR FILING DATE: 1998-09-18  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1225  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (491)..(1144)  
; OTHER INFORMATION:  
US-09-398-412B-1

Query Match 84.7%; Score 14.4; DB 4; Length 1225;  
Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGATAAAT 16  
|||||  
Db 277 TGTGGAAGGATAAAT 262

RESULT 9  
US-09-734-675-3  
; Sequence 3, Application US/09734675  
; Patent No. 6365391  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO00862  
; CURRENT APPLICATION NUMBER: US/09/734,675  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 38844  
; TYPE: DNA  
; ORGANISM: Human  
US-09-734-675-3

Query Match 84.7%; Score 14.4; DB 4; Length 38844;  
Best Local Similarity 93.8%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGATAAAT 16  
|||||  
Db 30049 TGTGGAAGGATAAAT 30064

RESULT 10  
US-09-128-155-16/c  
; Sequence 16, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03

; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 152331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(152331)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16

Query Match 84.7%; Score 14.4; DB 3; Length 152331;  
Best Local Similarity 93.8%; Pred. No. 3.5e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGATAAAT 16  
|||||  
Db 105263 TGTGGAAGGATAAAT 105248

RESULT 11  
US-09-128-155-17  
; Sequence 17, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 176373  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(176373)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-17

Query Match 84.7%; Score 14.4; DB 3; Length 176373;  
Best Local Similarity 93.8%; Pred. No. 3.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGTGGAAGGATAAAT 16  
|||||  
Db 78207 TGTGGAAGGATAAAT 78222

RESULT 12  
US-09-495-050A-121  
; Sequence 121, Application US/09495050A  
; Patent No. 6492505  
; GENERAL INFORMATION:  
; APPLICANT: Roopa, Reddy  
; APPLICANT: Guegler, Karl, J.  
; APPLICANT: Au-Young, Janice  
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED PROTEINS  
; FILE REFERENCE: PA-0013 US  
; CURRENT APPLICATION NUMBER: US/09/495,050A  
; CURRENT FILING DATE: 2000-01-31

;; PRIOR APPLICATION NUMBER: 60/118,318  
;; PRIOR FILING DATE: February 1, 1999  
;; NUMBER OF SEQ ID NOS: 305  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 121  
;; LENGTH: 234  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: Incyte ID No. 6492505 1386485CT1  
US-09-495-050A-121

Query Match 82.4%; Score 14; DB 4; Length 234;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGTGAAGGAATAAT 16  
DB 171 TGTGAAGGAATAAT 184

RESULT 13  
US-08-392-546C-4  
;; Sequence 4, Application US/08392546C  
;; Patent No. 5874298  
;; GENERAL INFORMATION:  
;; APPLICANT: Johnson, Janice, Kral, Robert M. Jr., Krapcho, Karen  
;; TITLE OF INVENTION: Insecticidal Toxins from Bracon hebetor  
;; NUMBER OF SEQUENCES: 11  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MADSON & METCALF  
;; STREET: 950 FIRST INTERSTATE BUILDING  
;; STREET: 170 SOUTH MAIN STREET  
;; CITY: SALT LAKE CITY  
;; STATE: UTAH  
;; COUNTRY: USA  
;; ZIP: 84101  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM  
;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: WORDPERFECT 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/392,546C  
;; FILING DATE: 2/17/95  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: L. CRAIG METCALF  
;; REGISTRATION NUMBER: 31,398  
;; REFERENCE/DOCKET NUMBER: 1094.2.3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (801) 537-1700  
;; TELEFAX: (801) 537-1799  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1127 nucleic acid  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: cdna  
;; HYPOTHETICAL:  
;; ANTI-SENSE: yes  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM: Bracon hebetor  
;; INDIVIDUAL ISOLATE: 30 kDa toxin cdna  
US-08-392-546C-4

Query Match 82.4%; Score 14; DB 2; Length 1127;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGAAGGAATAA 14

Db 156 TGTGAAGGAATAA 169

RESULT 14  
US-08-956-171E-53  
;; Sequence 53, Application US/08956171E  
;; Patent No. 6593114  
;; GENERAL INFORMATION:  
;; APPLICANT: Charles Kunsch  
;; Patrick S. Dillon  
;; Craig A. Rosen  
;; Steven C. Barash  
;; Michael R. Pannon  
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
;; NUMBER OF SEQUENCES: 5256  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Humar Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSPOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/956,171E  
;; FILING DATE: 20-Oct-1997  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/009,861  
;; FILING DATE: January 5, 1996  
;; APPLICATION NUMBER: 08/781,986  
;; FILING DATE: January 3, 1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mark J. Hyman  
;; REGISTRATION NUMBER: 46,789  
;; REFERENCE/DOCKET NUMBER: PB248P1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (240) 314-1224  
;; TELEFAX: (301) 309-8439  
;; INFORMATION FOR SEQ ID NO: 53:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 16592 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
US-08-956-171E-53

Query Match 82.4%; Score 14; DB 4; Length 16592;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGAAGGAATAA 14  
DB 13302 TGTGAAGGAATAA 13315

RESULT 15  
US-09-301-665-3/C  
;; Sequence 3, Application US/09301665  
;; Patent No. 6207876  
;; GENERAL INFORMATION:  
;; APPLICANT: KELLEMS, RODNEY E.  
;; APPLICANT: DATTA, SURJIT K.  
;; APPLICANT: BLACKBURN, MICHAEL R.  
;; TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND  
;; TITLE OF INVENTION: METHODS FOR THE USE THEREOF

```

; FILE REFERENCE: UTSH:243
; CURRENT APPLICATION NUMBER: US/09/301,665
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 60/083,408
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: 60/083,370
; EARLIER FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 36741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-301-665-3

Query Match      82.4%; Score 14; DB 3; Length 36741;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 TGAAGGATAAAT 16
Db      35748 TGAAGGATAAAT 35735

```

Search completed: March 25, 2004, 15:34:21  
Job time : 4.34382 secs

Result No.	Query %		DB	ID	Description	
	Score	Match Length				
1	25	100.0	6021	6	ABK48986 Genomic D	
2	100.0	6458	6	ABK48984	ABK48984 Genomic D	
3	19.8	79.2	1019	7	AD872007	AD872007 Rice gene
4	19.2	76.8	110000	7	ACHQ4281_1	Continuation (2 of
5	18.8	75.2	455	8	ACHQ39191	ACH39191 Human foe
6	18.8	75.2	464	8	ACHQ40492	ACH40492 Human foe
C 7	18.8	75.2	50000	4	AAF54867	Aaf54867 Nucleotid
C 8	18.8	75.2	81145	4	AAFS4868	AAF54868 Genomic n
C 9	18.8	75.2	494	7	ABK96482	ABK96482 Rice endo
C 10	18.4	73.6	816	7	AC34306	AC34306 Prokaryot
C 11	18.4	73.6	6347	4	ABL22340	ABL22340 Drosophil
12	18.4	73.6	110000	2	AAT42063_09	Continuation (10 o
13	18.4	73.6	110000	2	AAT42063_10	Continuation (11 o
C 14	18.2	72.8	686	6	ABS77078	ABS77078 Frog embr
15	18.2	72.8	5124	4	AAK75882	AAK75882 Human imm
C 16	18.2	72.8	32183	4	AAK35594	AAK35594 Human car
C 17	18.2	72.8	32183	9	AD846688	AD846688 Human car
C 18	18.2	72.8	51935	4	AAK75883	AAK75883 Human imm
19	18.2	72.8	349881	9	ADC86642	ADC86642 Human GPC
C 20	18.2	72.8	349901	9	ADC86940	ADC86940 Human GPC
C 21	18.2	72.8	349981	9	ADC87619	ADC87619 Human GPC
C 22	17.8	71.2	1119	5	AAS67702	AAS67702 DNA encod
C 23	17.8	71.2	1287	3	AAC42094	AAC42094 Arabidops

CC The invention describes an isolated human FOXC2 promoter region. The  
 CC mammalian FOXC2 promoter region for the modulation of FOXC2 polypeptide  
 CC expression has therapeutic value in treating type II diabetes mellitus,  
 CC obesity, hypercholesterolaemia, other cardiovascular diseases or  
 CC dyslipidaemias. This sequence encodes the mouse mesenchyme forkhead 1  
 CC (MHF-1) protein (also called FOXC2 transcription factor), described in  
 CC the method of the invention  
 XX  
 SQ Sequence 6021 BP; 1350 A; 1675 C; 1642 G; 1354 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 25; DB 6; Length 6021;  
 Best Local Similarity 100.0%; Pred. No. 0.79;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGTCCACAAATAAACAACTGTCC 25  
 DB 358 CAGTCCACAAATAAACAACTGTCC 382  
 RESULT 2  
 ABK48984  
 ID ABK48984 standard; DNA; 6458 BP.  
 AC ABK48984;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Genomic DNA encoding human transcription factor FOXC2.  
 KW Transcription factor; FOXC2; antidiabetic; anorectic; antilipaeimic;  
 KW cardiovascular; FOXC2 polypeptide expression modulator; FOXC2 promoter;  
 KW diabetes mellitus; obesity; hypercholesterolaemia; dyslipidaemia;  
 KW cardiovascular disease; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT exon 1..215  
 FT /tag= b  
 FT /note= "First exon according to the alternative  
 FT transcript"  
 FT 1..186  
 FT /tag= a  
 FT /tag= 187..215  
 FT /tag= c  
 FT /note= "Region coding for 5' part of alternative protein"  
 FT 215..216  
 FT /tag= d  
 FT /note= "Alternative first exon splice site"  
 FT 216..475  
 FT /tag= e  
 FT /note= "Fragment of the FOXC2 enhancer. Specifically  
 FT claimed in Claim 16"  
 FT 223..231  
 FT /tag= f  
 FT /note= "Fragment of the FOXC2 enhancer. Specifically  
 FT claimed in Claim 12"  
 FT 359..375  
 FT /tag= g  
 FT /note= "Fragment of the FOXC2 enhancer. Specifically  
 FT claimed in Claim 13"  
 FT 378..402  
 FT /tag= h  
 FT /note= "Fragment of the FOXC2 enhancer. Specifically  
 FT claimed in Claim 14"  
 FT 403..423  
 FT /tag= i  
 FT /note= "Fragment of the FOXC2 enhancer. Specifically  
 FT claimed in Claim 15"  
 FT 1250..2235  
 FT /tag= k  
 FT /note= "Fragment of the FOXC2 promoter. Specifically  
 FT claimed in Claim 3"

FT misc\_difference 1250..1749  
 FT /tag= j  
 FT /note= "Fragment of the FOXC2 promoter. Specifically  
 FT claimed in claim 2"  
 FT 1692..1703  
 FT misc\_difference  
 FT /tag= l  
 FT /note= "Fragment of the FOXC2 promoter. Specifically  
 FT claimed in claim 1"  
 FT 1746..4629  
 FT /tag= n  
 FT /note= "First exon according to the published form of the  
 FT transcript"  
 FT 1746..2234  
 FT /tag= m  
 FT 2235..3740  
 FT /tag= o  
 FT /product= "FOXC2"  
 FT /note= "Transcription factor"  
 FT 2448..2735  
 FT /tag= p  
 FT /note= "Region coding for DNA-binding domain"  
 FT 2516..4629  
 FT /tag= r  
 FT /note= "Second exon according to the alternative  
 FT transcript"  
 FT 2516..3740  
 FT /tag= q  
 FT /note= "Portion of polypeptide used in alternative  
 FT transcript"  
 FT 3741..4629  
 FT /tag= s  
 XX  
 PN WO200227008-A1.  
 XX  
 PD 04-APR-2002.  
 XX  
 XX 26-SEP-2001; 2001WO-SE002098.  
 PF  
 XX 26-SEP-2000; 2000SE-00003435.  
 PR  
 PR 10-OCT-2000; 2000US-0238897P.  
 PR 09-NOV-2000; 2000SE-00004102.  
 XX  
 PA (BIOV-) BIOVITRUM AB.  
 XX  
 PI Enerbaeck S, Krook K, Rondahl L, Wasserman WW;  
 XX  
 DR WPI; 2002-352129/38.  
 DR P-PSDB; AAU79816.  
 XX  
 PT An isolated FOXC2 promoter region that modulates the expression of a  
 PT FOXC2 polypeptide is useful for treating obesity and type II diabetes  
 PT mellitus.  
 XX  
 PS Claim 3; Page 33-38; 62pp; English.  
 XX  
 CC The invention describes an isolated human FOXC2 promoter region. The  
 CC mammalian FOXC2 promoter region for the modulation of FOXC2 polypeptide  
 CC expression has therapeutic value in treating type II diabetes mellitus,  
 CC obesity, hypercholesterolaemia, other cardiovascular diseases or  
 CC dyslipidaemias. This sequence encodes the human transcription factor  
 CC FOXC2, the promoter of which is described in the invention  
 XX  
 SQ Sequence 6458 BP; 1443 A; 1826 C; 1715 G; 1474 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 25; DB 6; Length 6458;  
 Best Local Similarity 100.0%; Pred. No. 0.79;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGTCCACAAATAAACAACTGTCC 25  
 DB 378 CAGTCCACAAATAAACAACTGTCC 402

```
RESULT 3
ADA72007
ID ADA72007 standard; DNA; 1019 BP.
XX
AC ADA72007;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5332.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX
XX Claim 27; SEQ ID NO 5332; 899bp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
XX Sequence 1019 BP; 298 A; 224 C; 263 G; 229 T; 0 U; 5 Other;
XX
XX Query Match 79.2%; Score 19.8; DB 7; Length 1019;
XX Best Local Similarity 91.3%; Pred. No. 98;
XX Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGTCCACAAATAAACAACACTGTGC 24
Db 657 AGTCCACAAATAAACAACACTGTGC 679

RESULT 4
ABQ84281.1
Continuation (2 of 5) of ABQ84281 from base 100001 (Chromosome 2 Asthma Locus DNA sequen-
WP Sequence split into 5 fragments LOCUS ABQ84281 Accession Abq84281
WP Fragment Name Begin End
WP ABQ84281_0 1 110000
WP ABQ84281_1 100001 210000
WP ABQ84281_2 200001 310000
WP ABQ84281_3 300001 410000
WP ABQ84281_4 400001 462586
XX
XX Query Match 76.8%; Score 19.2; DB 7; Length 110000;
XX Best Local Similarity 87.5%; Pred. No. 2.5e+02;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 AGTCCACAAATAAACAACACTGTGC 25
Db 13970 ATTCCACAAATATACAAACTATCC 13993

RESULT 5
ACH39191
ID ACH39191 standard; cDNA; 455 BP.
XX
XX ACH39191;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human foetal brain cDNA #558.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 26403; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030075623
XX
XX Sequence 455 BP; 134 A; 96 C; 102 G; 122 T; 0 U; 1 Other;
XX
XX Query Match 75.2%; Score 18.8; DB 8; Length 455;
XX Best Local Similarity 90.9%; Pred. No. 2.4e+02;
XX Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACACTGTGC 25
Db 376 TACACAAAAAACAACACTGTGC 397
```

## RESULT 6

ACH40492  
ID ACH40492 standard; cDNA; 464 BP.XX AC  
XX AC ACH40492;

XX DT 13-OCT-2003 (first entry)

XX DE Human foetal brain cDNA #1859.

XX OS Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

XX PA (DRMA/) DRMANAC R T.

XX PA (LABA/) LABAT I.

XX PA (STAC/) STACHE-CRAIN B.

XX PA (DICK/) DICKSON M C.

XX PA (JONE/) JONES L W.

XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 27704; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 464 BP; 135 A; 98 C; 109 G; 116 T; 0 U; 6 Other;

Query Match 75.2%; Score 18.8; DB 8; Length 464;

Best Local Similarity 90.9%; Pred. No. 2.4e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACTGTCC 25

DB 391 TACACAAAAAACAACTGTCC 412

## RESULT 7

AAF54867/c

ID AAF54867 standard; DNA; 50000 BP.

XX AC  
XX AC AAF54867;

XX DT 15-MAY-2001 (first entry)

XX DE Nucleotide sequence of a human transmembrane protein.

XX KW Human; transmembrane protein; 65h2 protein; 593 protein; prostaglandin;  
KW thromboxane; KIAA0880 protein; ss.

XX OS Homo sapiens.

XX PN WO200109185-A2.

XX PD 08-FEB-2001.

XX PF 28-JUL-2000; 2000WO-US020521.

XX PR 30-JUL-1999; 99US-00365162.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Curtis RA;

XX WPI; 2001-138648/14.

XX New nucleic acid molecules encoding transmembrane proteins designated  
PT 65h2 and 293 are useful for screening assays, detection assays and in  
PT predictive medicine.

XX Disclosure; Page 181-194; 215pp; English.

XX The present sequence encodes a human transmembrane protein. The  
CC specification describes transmembrane proteins designated KIAA0880, 65h2  
CC and 593. The proteins and polynucleotides can be used for screening  
CC assays, detection assays, e.g. chromosome mapping, tissue typing,  
CC forensic biology and predictive medicine, e.g. diagnostic assays,  
CC prognostic assays, monitoring clinical trials and pharmacogenomics. They  
CC polynucleotides can also be used for the recombinant production of 65h2  
CC and 593 proteins which can be used in assays to detect compounds which  
CC modulate their activity. These compounds are useful in the treatment of  
CC diseases characterized by aberrant activity or expression of protein  
CC which catalyse or facilitate transport of charged organic compound, e.g.  
CC prostaglandins and thromboxanes

XX Sequence 50000 BP; 13944 A; 11451 C; 11177 G; 13428 T; 0 U; 0 Other;

Query Match 75.2%; Score 18.8; DB 4; Length 50000;

Best Local Similarity 90.9%; Pred. No. 3.5e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACTGTCC 25

DB 1142 TACACAAAAAACAACTGTCC 1121

## RESULT 8

AAF54868/c

ID AAF54868 standard; DNA; 81145 BP.

XX AC  
XX AC AAF54868;

XX DT 15-MAY-2001 (first entry)

XX DE Genomic nucleotide sequence of a human 65h2 protein.

XX KW Human; transmembrane protein; 65h2 protein; 593 protein; prostaglandin;  
KW thromboxane; KIAA0880 protein; ss.

XX OS Homo sapiens.

XX PN WO200109185-A2.



XX PD 08-FEB-2001.  
XX PF  
XX PR 28-JUL-2000; 2000WO-US020521.  
XX PR 30-JUL-1999; 99US-00365162.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Curtis RA;  
XX DR WPI; 2001-138648/14.  
XX PT New nucleic acid molecules encoding transmembrane proteins designated  
XX PT 65h2 and 293 are useful for screening assays, detection assays and in  
XX PT predictive medicine.  
XX PS Disclosure; Fig 1L1-1L36; 215pp; English.  
XX CC The present sequence encodes a human transmembrane proteins designated  
XX CC 65h2. The specification also describes a transmembrane protein designated  
XX CC 593 and KIAA0880. The proteins and polynucleotides can be used for  
XX CC screening assays, detection assays, e.g. chromosome mapping, tissue  
XX CC typing, forensic biology and predictive medicine, e.g. diagnostic assays,  
XX CC prognostic assays, monitoring clinical trials and pharmacogenomics. They  
XX CC polynucleotides can also be used for the recombinant production of 65h2  
XX CC and 593 proteins, which can be used in assays to detect compounds which  
XX CC modulate their activity. These compounds are useful in the treatment of  
XX CC diseases characterized by aberrant activity or expression of protein  
XX CC which catalyze or facilitate transport of charged organic compound, e.g.  
XX CC prostaglandins and thromboxanes  
XX SQ Sequence 81145 BP; 22396 A; 18624 C; 18484 G; 21641 T; 0 U; 0 Other;  
Query Match 75.2%; Score 18.8; DB 4; Length 81145;  
Best Local Similarity 90.9%; Pred. No. 3.6e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 TCCACAAATAAACAACACTGTC 25  
DB 1142 TACACAAAATAAACAACACTGTC 1121  
RESULT 9  
ID ABX96482 standard; cDNA; 494 BP.  
XX AC ABX96482;  
XX DT 13-MAY-2003 (first entry)  
XX DE Rice endosperm expression sequence label #23.  
XX KW Rice; endosperm expression sequence label; gene chip; babel technique;  
XX KW microarray; ss.  
XX OS Oryza sp.  
XX PN CN1366050-A.  
XX PD 28-AUG-2002.  
XX PF 31-OCT-2001; 2001CN-00135874.  
XX PR 31-OCT-2001; 2001CN-00135874.  
XX PA (UYZH-) UNIV ZHEJIANG.  
XX PI Dong H, Li D;  
XX DR WPI; 2003-240400/24.  
XX PT Rice endosperm expression sequence label and genechip prepared from it.

XX PS Claim 1; Page 18 (Disclosure); 29pp; Chinese.  
XX CC The invention describes a novel rice endosperm expression sequence label  
XX CC and the gene chip formed from it. The expression sequence babel technique  
XX CC is used to create a rice endosperm cDNA library. Non-redundant expression  
XX CC sequence labels are then combined and used to create a gene chip by  
XX CC microarray techniques. This sequence represents a rice endosperm  
XX CC expression sequence  
XX SQ Sequence 494 BP; 132 A; 117 C; 102 G; 143 T; 0 U; 0 Other;  
Query Match 74.4%; Score 18.6; DB 7; Length 494;  
Best Local Similarity 84.0%; Pred. No. 2.9e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 CAGTCCACAAATAAACAACACTGTC 25  
DB 252 CAATACACAAATAAATAAAGTCCC 276  
RESULT 10  
AC34306/c  
ID ACA34306 standard; DNA; 816 BP.  
XX AC ACA34306;  
XX DT 19-JUN-2003 (first entry)  
XX DE Prokaryotic essential gene #15963.  
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX KW drug design; gene.  
XX OS Haemophilus influenzae.  
XX PN WC200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
XX DR P-PSDB; ABU30436.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 14; SEQ ID NO 22176; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a gene required for proliferation or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 816 BP; 255 A; 137 C; 172 G; 252 T; 0 U; 0 Other;  
Query Match 73.6%; Score 18.4; DB 7; Length 816;  
Best Local Similarity 95.0%; Pred. No. 3.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CCACAAATAAACAACACTGTC 24

Db 539 CCACAAATAAACAACACTTC 520

RESULT 11

ABL22340/c  
ID ABL22340 standard; DNA; 6347 BP.

XX AC ABL22340;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 18493.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX RW pharmaceutical; gene; ds.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from *Drosophila* and for elucidating cell signaling and cell-cell  
XX FT interactions.

XX PS Claim 1; SEQ ID NO 18493; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 6347 BP; 1734 A; 1346 C; 1299 G; 1968 T; 0 U; 0 Other;

Query Match 73.6%; Score 18.4; DB 4; Length 6347;  
Best Local Similarity 95.0%; Pred. No. 4.3e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAGTCCCAATAAACAAC 20

Db 5626 CAGTCCCAATAAACAAC 5607

RESULT 12

AAT42063\_09

Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete g  
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

Fragment Name	Begin	End
WP AAT42063_00	1	110000
WP AAT42063_01	100001	210000
WP AAT42063_02	200001	310000
WP AAT42063_03	300001	410000
WP AAT42063_04	400001	510000
WP AAT42063_05	500001	610000
WP AAT42063_06	600001	710000
WP AAT42063_07	700001	810000
WP AAT42063_08	800001	910000
WP AAT42063_09	900001	1010000
WP AAT42063_10	1000001	1110000
WP AAT42063_11	1100001	1210000
WP AAT42063_12	1200001	1310000
WP AAT42063_13	1300001	1410000
WP AAT42063_14	1400001	1510000
WP AAT42063_15	1500001	1610000
WP AAT42063_16	1600001	1710000
WP AAT42063_17	1700001	1810000
WP AAT42063_18	1800001	1830121

Query Match 73.6%; Score 18.4; DB 2; Length 110000;  
Best Local Similarity 95.0%; Pred. No. 5.4e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CCACAAATAAACAACACTGTC 24

Db 105260 CCACAAATAAACAACACTTC 105279

RESULT 13

AAT42063\_10

Continuation (11 of 19) of AAT42063 from base 1000001 (Haemophilus influenzae complete g  
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

Fragment Name	Begin	End
WP AAT42063_00	1	110000
WP AAT42063_01	100001	210000
WP AAT42063_02	200001	310000
WP AAT42063_03	300001	410000
WP AAT42063_04	400001	510000
WP AAT42063_05	500001	610000
WP AAT42063_06	600001	710000
WP AAT42063_07	700001	810000
WP AAT42063_08	800001	910000
WP AAT42063_09	900001	1010000
WP AAT42063_10	1000001	1110000
WP AAT42063_11	1100001	1210000
WP AAT42063_12	1200001	1310000
WP AAT42063_13	1300001	1410000
WP AAT42063_14	1400001	1510000
WP AAT42063_15	1500001	1610000
WP AAT42063_16	1600001	1710000

WP AAT42063\_17 1700001 1810000  
WP AAT42063\_18 1800001 1830121

Query Match 73.6%; Score 18.4; DB 2; Length 110000;  
Best Local Similarity 95.0%; Pred. No. 5.4e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CCACAAATAAACAACACTGTC 24  
Dd 5260 CCACAAATAAACAACACTTC 5279

RESULT 14  
ABS77078/c  
ID ABS77078 standard; cDNA; 686 BP.  
XX  
AC ABS77078;  
XX  
DT 12-DEC-2002 (first entry)  
XX  
DE Frog embryonic gene sequence Q9925471.  
XX  
KW Frog; ss; embryonic development; developmental disorder; microarray;  
KW cell differentiation.  
XX  
OS Xenopus laevis.  
XX  
PN US2002081610-A1.  
XX  
PD 27-JUN-2002.  
XX  
PF 23-JUL-2001; 2001US-00910943.  
XX  
PR 21-JUL-2000; 2000US-02196568P.  
XX  
PA (UVRQ ) UNIV ROCKEFELLER.  
XX

PI Hemmati-Brivanlou A, Altmann CR;  
XX  
DR WPI; 2002-626534/67.  
XX  
PT Nucleic acid array containing Xenopus embryonic nucleic acids is useful  
PT to identify genes involved in embryonic development, to identify  
PT different types of embryonic cells, and to diagnose developmental  
PT disorders.  
XX  
PS Claim 1; Page 348; 823pp; English.  
XX

CC The invention relates to a nucleic acid array, where each coordinate  
CC contains a single nucleic acid species having one of 770 nucleotide  
CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene  
CC product, or its complement or hybridisable fragment of not less than 20  
CC contiguous nucleotides of one of those sequences. Also included are  
CC detecting differential expression of embryonic genes, comprising: (a)  
CC contacting a nucleic acid array comprising genes expressed in embryonic  
CC but not mature cells with nucleic acids from sample and control cells;  
CC and (b) detecting differential hybridisation of nucleic acids from the  
CC sample cells relative to the control cells; and detecting defects in  
CC development, comprising: (a) contacting nucleic acids from test cells  
CC undergoing development with a nucleic acid array of gene products known  
CC to play a fundamental role in the development process; and (b) detecting  
CC a difference in expression of a fundamental gene in the sample cells  
CC relative to a standard. The invention is useful to identify genes  
CC involved in embryonic development and related processes such as cell  
CC differentiation. This would be useful for diagnosing developmental  
CC disorders and for identifying different types of embryonic cells. The  
CC present sequence is one of the 770 Xenopus embryonic cDNA sequences  
XX  
SQ Sequence 686 BP; 187 A; 144 C; 170 G; 184 T; 0 U; 1 Other;

Query Match 72.8%; Score 18.2; DB 6; Length 686;  
Best Local Similarity 87.0%; Pred. No. 4.4e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACACTGT 23  
Dd 345 CAGTCCACAAATGACCAACCGT 323

RESULT 15  
AAK75882  
ID AAK75882 standard; DNA; 5124 BP.  
XX  
AC AAK75882;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30694.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-02311344P.  
PR 08-SEP-2000; 2000US-02311413P.  
PR 08-SEP-2000; 2000US-02311414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-02311968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 12-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 01-DEC-2000; 2000US-0250392P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 06-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251889P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX  
PS Disclosure; SEQ ID NO 30694; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to prevent the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK4703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 5124 BP; 1322 A; 1253 C; 1083 G; 1466 T; 0 U; 0 Other;  
Query Match 72.8%; Score 18.2; DB 4; Length 5124;  
Best Local Similarity 87.0%; Pred. No. 5.2e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 GTCCACAAATAAACAACTGTCC 25  
DB 5025 GTCCACAGATCAACAATCTCC 5047  
Search completed: March 25, 2004, 10:25:20  
Job time : 29.9441 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 09:55:14 ; Search time 4.91739 Seconds  
(without alignments)  
2821.370 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_378\_402

Perfect score: 25  
Sequence: 1 cagtcacaaataaactgccc 25

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	73.6	1830121	4 US-09-557-884-1	Sequence 1, Appli
2	18.4	73.6	1830121	4 US-09-643-990A-1	Sequence 1, Appli
3	17.6	70.4	281	4 US-09-313-294A-195	Sequence 195, App
4	17.6	70.4	1056	4 US-09-543-681A-2690	Sequence 2690, App
5	17.6	70.4	3615	4 US-09-221-017B-971	Sequence 971, App
6	17.6	70.4	148567	4 US-09-801-876B-3	Sequence 3, Appli
7	17.6	70.4	148567	4 US-10-254-869-3	Sequence 3, Appli
8	17.4	69.6	350	3 US-09-157-177-110	Sequence 110, App
9	17.4	69.6	2411	4 US-09-446-301A-15	Sequence 15, Appl
10	17.4	69.6	2411	4 US-09-099-932-47	Sequence 47, Appl
11	17.2	68.8	2559	4 US-09-319-588C-17	Sequence 17, Appl
12	17.2	68.8	9183	4 US-09-319-588C-1	Sequence 1, Appli
13	17.2	68.8	1230025	4 US-09-198-452A-1	Sequence 1, Appli
14	17	68.0	438	4 US-09-540-236-476	Sequence 476, App
15	17	68.0	505	4 US-09-833-381-1109	Sequence 1109, Ap
16	16.6	66.4	1796	3 US-09-276-531-113	Sequence 113, App
17	16.6	66.4	2412	4 US-09-543-681A-2595	Sequence 2595, Ap
18	16.6	66.4	4223	3 US-09-541-782-5	Sequence 5, Appli
19	16.6	66.4	4223	4 US-09-723-820-5	Sequence 5, Appli
20	16.6	66.4	4223	4 US-10-270-085-5	Sequence 5, Appli
21	16.6	66.4	5220	4 US-09-540-236-1819	Sequence 1819, Ap
22	16.6	66.4	63563	4 US-09-596-002-33	Sequence 33, Appl
23	16.6	66.4	786431	4 US-09-751-389-3	Sequence 3, Appli
24	16.4	65.6	220	4 US-09-679-409-110	Sequence 110, App
25	16.4	65.6	714	4 US-09-679-409-109	Sequence 109, App
26	16.4	65.6	1158	4 US-09-539-333D-76	Sequence 76, Appl
27	16.4	65.6	319608	4 US-09-539-333D-1	Sequence 1, Appli

28 16.4 65.6 319608 4 US-09-679-409-1 Sequence 1, Appli  
29 16.2 64.8 395 4 US-09-319-056B-22 Sequence 22, Appl  
30 16.2 64.8 395 4 US-09-319-056B-24 Sequence 24, Appl  
31 16.2 64.8 759 4 US-09-134-000C-1426 Sequence 1426, Ap  
32 16.2 64.8 867 4 US-09-107-532A-2355 Sequence 2355, Ap  
33 16.2 64.8 1056 4 US-09-328-332-2030 Sequence 2030, Ap  
34 16.2 64.8 14113 3 US-09-223-134-1 Sequence 1, Appli  
35 16.2 64.8 14113 3 US-08-992-801-1 Sequence 1, Appli  
36 16.2 64.8 14113 3 US-09-223-535-1 Sequence 1, Appli  
37 16.2 64.8 1830121 4 US-09-557-884-1 Sequence 1, Appli  
38 16.2 64.8 1830121 4 US-09-643-990A-1 Sequence 1, Appli  
39 16 64.0 34 2 US-08-555-723B-11 Sequence 11, Appl  
40 16 64.0 34 3 US-09-123-485-11 Sequence 11, Appl  
41 16 64.0 42 2 US-08-343-443B-54 Sequence 54, Appl  
42 16 64.0 336 4 US-09-134-000C-2101 Sequence 2101, Ap  
43 16 64.0 585 4 US-09-134-000C-1803 Sequence 1803, Ap  
44 16 64.0 651 4 US-09-096-724B-1 Sequence 1, Appli  
45 16 64.0 662 4 US-08-576-202-32 Sequence 32, Appl

ALIGNMENTS

RESULT 1  
US-09-557-884-1  
; Sequence 1, Application US/09557884  
; Patent No. 6306581  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; Thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 1/2 inch diskette  
; COMPUTER: Dell Pentium  
; OPERATING SYSTEM: MS DOS v6.22  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/557,884  
; FILING DATE: 25-Apr-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/476,102  
; FILING DATE: JUN-5-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB186P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1830121 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: Double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Query Match 73.6%; Score 18.4; DB 4; Length 1830121;  
Best Local Similarity 95.0%; Pred. No. 76;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CCACAAATAAACAACTGTC 24

Db 1005260 CCACAAATAAACAACTTC 1005279

## RESULT 2

US-09-643-990A-1  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289

## GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann  
Mark D. Adams  
Owen White  
Hamilton O. Smith  
J. Craig Venter

TITLE OF INVENTION: the Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof

## NUMBER OF SEQUENCES: 1

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA

ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21

## ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

## SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 73.6%; Score 18.4; DB 4; Length 1830121;  
Best Local Similarity 95.0%; Pred. No. 76;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CCACAAATAAACAACTGTC 24

Db 1005260 CCACAAATAAACAACTTC 1005279

## RESULT 3

US-09-313-294A-195  
; Sequence 195, Application US/09313294A  
; Patent No. 6476212

## GENERAL INFORMATION:

APPLICANT: Lalgudi, Raghunath V.  
Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 195  
; LENGTH: 281  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700548714H1  
US-09-313-294A-195

Query Match 70.4%; Score 17.6; DB 4; Length 281;  
Best Local Similarity 83.3%; Pred. No. 65;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACTGTC 24

Db 206 CAGTCCAGATATAACACACTGAC 229

## RESULT 4

US-09-543-681A-2690  
; Sequence 2690, Application US/09543681A  
; Patent No. 6805709

## GENERAL INFORMATION:

APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709-1002-001

CURRENT APPLICATION NUMBER: US/09/543,681A

CURRENT FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 2690

LENGTH: 1056

TYPE: DNA

ORGANISM: Proteus mirabilis

US-09-543-681A-2690

Query Match 70.4%; Score 17.6; DB 4; Length 1056;  
Best Local Similarity 83.3%; Pred. No. 76;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACTGTC 24

Db 187 CAATCCACAGAGAAACAACTCTC 210

## RESULT 5

US-09-221-017B-971  
; Sequence 971, Application US/09221017B  
; Patent No. 6444799

## GENERAL INFORMATION:

APPLICANT: Ross, Bruce C.  
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
NUMBER OF SEQUENCES: 1120  
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 PAGE MILL ROAD

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

Query Match	Score 17.6;	DB 4;	Length 148567;
70.4%	Score 17.6;	DB 4;	Length 148567;

Mon Mar 29 10:42:09 2004

```
US-09-319-588C-17/c
; Sequence 17, Application US/09319588C
; Patent No. 6509018
; GENERAL INFORMATION:
; APPLICANT: INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE-INSERM
; APPLICANT: ASSISTANCE PUBLIQUE-HOPITAUX DE PARIS
; APPLICANT: INSTITUT PASTEUR
; APPLICANT: MAUCLERE, Philippe
; APPLICANT: LOUSSET-AJAKA, Ibtissam
; APPLICANT: SIMON, Francois
; APPLICANT: SARAGOSTI, Sentob
; APPLICANT: BARRE-SINOUSI, Francoise
; TITLE OF INVENTION: NON-M NON-O HIV STRAINS, FRAGMENTS AND APPLICATIONS.
; FILE REFERENCE: 598US12
; CURRENT APPLICATION NUMBER: US/09/319,588C
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: FR96/15087
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) (2556)
US-09-319-588C-17
Query Match 68.8%; Score 17.2; DB 4; Length 2559;
Best Local Similarity 86.4%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GTCCACAAATAACAACTGTC 24
|||||
DB 1247 GTCCACAAATTTACAATCTGTC 1226

RESULT 12
US-09-319-588C-1/c
; Sequence 1, Application US/09319588C
; Patent No. 6509018
; GENERAL INFORMATION:
; APPLICANT: INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE-INSERM
; APPLICANT: ASSISTANCE PUBLIQUE-HOPITAUX DE PARIS
; APPLICANT: INSTITUT PASTEUR
; APPLICANT: MAUCLERE, Philippe
; APPLICANT: LOUSSET-AJAKA, Ibtissam
; APPLICANT: SIMON, Francois
; APPLICANT: SARAGOSTI, Sentob
; APPLICANT: BARRE-SINOUSI, Francoise
; TITLE OF INVENTION: NON-M NON-O HIV STRAINS, FRAGMENTS AND APPLICATIONS.
; FILE REFERENCE: 598US12
; CURRENT APPLICATION NUMBER: US/09/319,588C
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: FR96/15087
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9183
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-319-588C-1
Query Match 68.8%; Score 17.2; DB 4; Length 9183;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GTCCACAAATAACAACTGTC 24
|||||
DB 7056 GTCCACAAATTTACAATCTGTC 7035

US-09-446-301A-15/c
; Sequence 15, Application US/09446301A
; Patent No. 6506893
; GENERAL INFORMATION:
; APPLICANT: EL SOLH, NEVINE
; APPLICANT: ALLIGNET, JEANINE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03715-0059
; CURRENT APPLICATION NUMBER: US/09/446,301A
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Staphylococcus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (700...2355, 2388..2411)
US-09-446-301A-15
Query Match 69.6%; Score 17.4; DB 4; Length 2411;
Best Local Similarity 94.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ACAATAAACAACTGTC 25
|||||
DB 575 ACAATAAACAACTATCC 557

RESULT 10
US-09-099-932-47/c
; Sequence 47, Application US/09099932
; Patent No. 6570001
; GENERAL INFORMATION:
; APPLICANT: El Solh, Nevine
; APPLICANT: Allignet, Jeanine
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03495-0173-00000
; CURRENT APPLICATION NUMBER: US/09/099,932
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,380
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 47
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Staphylococcus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (700)..(2355)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2388)..(2411)
US-09-099-932-47
Query Match 69.6%; Score 17.4; DB 4; Length 2411;
Best Local Similarity 94.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ACAATAAACAACTGTC 25
|||||
DB 575 ACAATAAACAACTATCC 557

RESULT 11
```



```

RESULT 13
US-09-198-452A-1
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
TITLE OF INVENTION:
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (165001)..(180000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (195001)..(210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (210001)..(225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (225001)..(240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (240001)..(255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (255001)..(270000)
OTHER INFORMATION: n=a or c or g or t
LOCATION: (270001)..(285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (285001)..(300000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t

```

LOCATION: (630001)...(645000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (645001)...(660000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (660001)...(675000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (675001)...(690000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (690001)...(705000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (705001)...(720000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (720001)...(735000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (735001)...(750000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (750001)...(765000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (765001)...(780000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (780001)...(795000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (795001)...(810000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (810001)...(825000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (825001)...(840000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (840001)...(855000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (855001)...(870000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (870001)...(885000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (885001)...(900000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (900001)...(915000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature

Query Match 68.8%; Score 17.2; DB 4; Length 1230025;  
Best Local Similarity 86.4%; Pred. No. 2e+02; 3; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0

Qy 4 TCACAAATAAACAACACTGTCC 25  
Db 447969 TCACAACTAACGAGCTGTCC 447990

RESULT 14  
US-09-540-236-476/c  
; Sequence 476, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 476  
LENGTH: 438  
TYPE: DNA  
ORGANISM: M.catarrhalis  
US-09-540-236-476

Query Match 68.0%; Score 17; DB 4; Length 438;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CAGTCCACAATAAACAACACTGTCC 25  
Db 192 CAGTCCCAATAACAACAACAAATGGCC 168

RESULT 15  
US-09-833-381-1109/c  
; Sequence 1109, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
FILE REFERENCE: 5800-119  
CURRENT APPLICATION NUMBER: US/09/833,381  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 09/516,448  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 2050  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1109  
LENGTH: 505  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(505)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1109

Query Match 68.0%; Score 17; DB 4; Length 505;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CAGTCCACAATAAACAACACTGTCC 25  
Db 335 CAGTCCCAATAACAACCTACTATGC 311

Search completed: March 25, 2004, 15:34:27  
JOB time : 10.9174 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 10:03:04 ; Search time 21.6562 Seconds

(without alignments)  
4297.861 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_378\_402

Perfect score: 25

Sequence: 1 cagtcacaaataaacaactgtcc 25

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504946 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCF\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCFUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	6021	9	US-09-963-285-5
2	25	100.0	6458	9	US-09-963-285-1
3	20.2	80.8	1907	15	US-10-027-632-53050
4	19.2	76.8	202802	15	US-10-085-117-286
5	18.8	75.2	455	10	US-09-918-995-26403
6	18.8	75.2	464	10	US-09-918-995-27704
7	18.8	75.2	5000	13	US-10-063-763-4
8	18.8	75.2	5000	14	US-10-060-763-4
9	18.6	74.4	437	15	US-10-027-632-2550
10	18.4	73.6	816	12	US-10-282-122A-22176
11	18.4	73.6	1830121	14	US-10-329-960-1
12	18.4	73.6	1830121	15	US-10-329-960-1
13	18.2	72.8	686	9	US-09-910-943-313
14	18.2	72.8	32183	9	US-09-764-869-1494
15	18.2	72.8	32183	14	US-10-091-504-1494

c	16	18.2	72.8	32183	15	US-10-227-577-1494	Sequence 1494, Ap
	17	18.2	72.8	358246	15	US-10-292-798-1095	Sequence 1095, Ap
	18	18.2	72.8	744802	15	US-10-292-798-1369	Sequence 1369, Ap
c	19	18.2	72.8	822800	15	US-10-292-798-1393	Sequence 1393, Ap
	20	18	72.0	1087	15	US-10-027-632-118026	Sequence 118026, Ap
c	21	17.8	71.2	9007	12	US-10-221-7144-337	Sequence 337, App
	22	17.6	70.4	406	15	US-10-027-632-64593	Sequence 64593, A
	23	17.6	70.4	447	15	US-10-027-632-310659	Sequence 310659, A
	24	17.6	70.4	452	15	US-10-027-632-36169	Sequence 36169, A
	25	17.6	70.4	557	15	US-10-027-632-51602	Sequence 51602, A
	26	17.6	70.4	557	15	US-10-027-632-298785	Sequence 298785, A
	27	17.6	70.4	557	15	US-10-027-632-321555	Sequence 321555, A
	28	17.6	70.4	595	15	US-10-027-632-259330	Sequence 259330, A
	29	17.6	70.4	595	15	US-10-027-632-259331	Sequence 259331, A
	30	17.6	70.4	595	15	US-10-027-632-259332	Sequence 259332, A
	31	17.6	70.4	1026	12	US-10-282-122A-32601	Sequence 32601, A
	32	17.6	70.4	1434	12	US-10-425-114-3242	Sequence 3242, Ap
	33	17.6	70.4	1702	12	US-10-425-114-13532	Sequence 13532, A
	34	17.6	70.4	5514	9	US-09-966-880A-9	Sequence 9, Appli
c	35	17.6	70.4	8605	14	US-10-311-455-507	Sequence 507, App
	36	17.6	70.4	11204	9	US-09-966-880A-35	Sequence 35, Appl
c	37	17.6	70.4	16167	14	US-10-311-455-1055	Sequence 1055, Ap
	38	17.6	70.4	16167	14	US-10-240-485-81	Sequence 81, Appl
c	39	17.6	70.4	148567	9	US-09-801-876B-3	Sequence 3, Appli
	40	17.6	70.4	148567	12	US-10-657-442-3	Sequence 3, Appli
	41	17.6	70.4	148567	14	US-10-254-869-3	Sequence 3, Appli
	42	17.4	69.6	1120	9	US-09-938-842A-2767	Sequence 2767, Ap
	43	17.4	69.6	1120	11	US-09-938-842A-2767	Sequence 2767, Ap
c	44	17.4	69.6	1290	15	US-10-398-321-3094	Sequence 3094, Ap
c	45	17.4	69.6	2411	14	US-10-253-904-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-09-963-285-5  
; Sequence 5, Application US/09963285  
; Patent No. US30020090707A1  
; GENERAL INFORMATION:  
; APPLICANT: Emerck, Sven  
; APPLICANT: Krook, Katrina  
; APPLICANT: Rondahl, Lena  
; APPLICANT: Wasserman, Wyeth  
; TITLE OF INVENTION: PROMOTER SEQUENCES  
; FILE REFERENCE: 13425-042001  
; CURRENT APPLICATION NUMBER: US/09/963,285  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: SE 0004102-0  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/238,897  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: SE 0003435-5  
; PRIOR FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 6021  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1649)...(439)  
US-09-963-285-5

Query Match 100.0%; Score 25; DB 9; Length 6021;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CAGTCCACAAATAAACAACACTGTCC 25  
Db 358 CAGTCCACAAATAAACAACACTGTCC 382

```

; LOCATION: (1)...(1907)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-53050

Query Match      80.8%; Score 20.2; DB 15; Length 1907;
Best Local Similarity 88.0%; Pred. No. 62;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAGTCCACAAATAAACAAACTGTCC 25
Db 283 CATTCCACAGATAAACAAACCGTCC 307

RESULT 4
US-10-085-117-286
; Sequence 286, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 202802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-286

Query Match      76.8%; Score 19.2; DB 15; Length 202802;
Best Local Similarity 87.5%; Pred. No. 4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAGTCCACAAATAAACAAACTGTCTC 24
Db 176350 CAGACCAACAAATAAACAAACCCCTC 176373

RESULT 5
US-09-918-995-26403
; Sequence 26403, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26403
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26403

Query Match      75.2%; Score 18.8; DB 10; Length 455;
Best Local Similarity 90.9%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TCACAAATAAACAAACTGTCC 25

```

```

; Sequence 1, Application US/09963285
; Patent No. US20020090707A1
; GENERAL INFORMATION:
; APPLICANT: Emerbeck, Sven
; APPLICANT: Krook, Katarina
; APPLICANT: Rondahl, Lena
; APPLICANT: Wasserman, Wyeth
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-042001
; CURRENT APPLICATION NUMBER: US/09/963,285
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: SE 0004102-0
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/238,897
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: SE 0003435-5
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2235)...(3737)
US-09-963-285-1

Query Match      100.0%; Score 25; DB 9; Length 6458;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGTCCACAAATAAACAAACTGTCC 25
Db 378 CAGTCCACAAATAAACAAACTGTCC 402

RESULT 3
US-10-027-632-53050
; Sequence 53050, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53050
; LENGTH: 1907
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```

```
Db 376 TACACAAAAAACAACACTGTCC 397
| ||||| ||||| ||||| ||||| |||||
; TITLE OF INVENTION: NO. US20030022286A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/060,763
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-763-4

Query Match 75.2%; Score 18.8; DB 14; Length 50000;
Best Local Similarity 90.9%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACACTGTCC 25
| ||||| ||||| ||||| ||||| |||||
Db 1142 TACACAAAAAACAACACTGTCC 1121

RESULT 9
US-10-027-632-2550
; Sequence 2550, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2550
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-2550

Query Match 74.4%; Score 18.6; DB 15; Length 437;
Best Local Similarity 84.0%; Pred. No. 2.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACACTGTCC 25
| ||||| ||||| ||||| ||||| |||||
Db 235 CAGTCCACAAATAAACAACACTGTCC 259

RESULT 10
US-10-282-122A-22176/c
; Sequence 22176, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl

Db 376 TACACAAAAAACAACACTGTCC 397
| ||||| ||||| ||||| ||||| |||||
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27704
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(464)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27704

Query Match 75.2%; Score 18.8; DB 10; Length 464;
Best Local Similarity 90.9%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACACTGTCC 25
| ||||| ||||| ||||| ||||| |||||
Db 391 TACACAAAAAACAACACTGTCC 412

RESULT 7
US-10-063-763-4/c
; Sequence 4, Application US/10063763
; Publication No. US20010001663A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: NO. US20010001663A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/063,763
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 09/365,162
; PRIOR FILING DATE: 1999-JUL-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-063-763-4

Query Match 75.2%; Score 18.8; DB 13; Length 50000;
Best Local Similarity 90.9%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCCACAAATAAACAACACTGTCC 25
| ||||| ||||| ||||| ||||| |||||
Db 1142 TACACAAAAAACAACACTGTCC 1121

RESULT 8
US-10-060-763-4/c
; Sequence 4, Application US/10060763
; Publication No. US20030022286A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
```

```

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22176
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; US-10-282-122A-22176

Query Match 73.6%; Score 18.4; DB 12; Length 816;
Best Local Similarity 95.0%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CCACAAATATAACAACCTGTC 24
Db 539 CCACAAATATAACAACCTTC 520

RESULT 11
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/543,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```

```
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
```

```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

```
Query Match 73.6%; Score 18.4; DB 14; Length 1830121;
Best Local Similarity 95.0%; Pred. No. 9.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 5 CCACAAATAAACAAACTGTC 24
|||||||
Db 1005260 CCACAAATAAACAAACTTC 1005279
```

```
RESULT 12
US-10-329-670-1
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10/329,670
CURRENT FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
```

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (40808)..(40810)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (44975)..(44975)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (45593)..(45593)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (45732)..(45732)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (47036)..(47036)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (51334)..(51334)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (51602)..(51602)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (51786)..(51786)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (51805)..(51805)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (55369)..(55369)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (65309)..(65309)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:

NAME/KEY: misc feature  
LOCATION: (65313)..(65313)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (80024)..(80024)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (100091)..(100091)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (102696)..(102696)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (105121)..(105121)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (107248)..(107248)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (117136)..(117136)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (119750)..(119750)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (119924)..(119924)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (120038)..(120038)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (121344)..(121344)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122167)..(122167)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122336)..(122336)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (131340)..(131340)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (131360)..(131360)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (139910)..(139910)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (140398)..(140398)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (142750)..(142750)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature



LOCATION: (145058)..(145058)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (145171)..(145171)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (145942)..(145942)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (147197)..(147197)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (150841)..(150841)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (152500)..(152500)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (152530)..(152530)

Query Match 73.6%; Score 18.4; DB 15; Length 1830121;  
Best Local Similarity 95.0%; Pred. No. 9.9e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CCACAAATAAACAACACTGTC 24  
|||||  
Db 1005260 CCACAAATAAACAACACTTTC 1005279

RESULT 13  
US-09-910-943-313/c  
Sequence 313, Application US/09910943  
Patent No. US20020081610A1  
GENERAL INFORMATION:  
APPLICANT: Hemmati-Brihanlou, Ali  
APPLICANT: Altman, Curtis  
TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression  
FILE REFERENCE: 7529/1G148US1  
CURRENT APPLICATION NUMBER: US/09/910,943  
CURRENT FILING DATE: 2001-07-23  
NUMBER OF SEQ ID NOS: 742  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 313  
LENGTH: 686  
TYPE: DNA  
ORGANISM: Xenopus laevis  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(686)  
OTHER INFORMATION: n may be a or g or c or t/u  
US-09-910-943-313

Query Match 72.8%; Score 18.2; DB 9; Length 686;  
Best Local Similarity 87.0%; Pred. No. 3.7e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAGTCCACAAATAAACAACACTGT 23  
|||||  
Db 345 CAGTCCACAAATGACCAACGGT 323

RESULT 14  
US-09-764-869-1494/c  
Sequence 1494, Application US/09764869  
Patent No. US20020061521A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC007  
CURRENT APPLICATION NUMBER: US/09/764,869  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2442  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1494  
LENGTH: 32183  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-869-1494

Query Match 72.8%; Score 18.2; DB 9; Length 32183;  
Best Local Similarity 87.0%; Pred. No. 7.7e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GTCCACAAATAAACAACACTGTCC 25  
|||||  
Db 14935 GTCCACAGATCAACAACACTCTCC 14913

RESULT 15  
US-10-091-504-1494/c  
Sequence 1494, Application US/10091504  
Publication No. US2003005908A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC007C1  
CURRENT APPLICATION NUMBER: US/10/091,504  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 2442  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1494  
LENGTH: 32183  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-091-504-1494

Query Match 72.8%; Score 18.2; DB 14; Length 32183;  
Best Local Similarity 87.0%; Pred. No. 7.7e+02;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GTCCACAAATAAACAACACTGTCC 25  
|||||  
Db 14935 GTCCACAGATCAACAACACTCTCC 14913

Search completed: March 25, 2004, 15:52:54  
Job time : 26.6562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 08:39:03 ; Search time 170.338 Seconds  
(without alignments)  
6361.316 Million cell updates/sec

Title: US-09-963-285-1\_COPY\_378\_402

Perfect score: 25

Sequence: 1 cagtcacacataaacaactgcc 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.hcg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.ats.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.hcg.hum.\*

31: em.hcg.inv.\*

32: em.hcg.other.\*

33: em.hcg.mus.\*

34: em.hcg.pln.\*

35: em.hcg.rtd.\*

36: em.hcg.ham.\*

37: em.hcg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	100.0	6021	10	MMFHEAD1	Y08222 M.musculus
2	25	100.0	168656	9	AC009108	AC009108 Homo sapi
3	25	100.0	178416	10	AC127554	AC127554 Mus muscu
4	21.4	85.6	162337	2	AC101961	AC101961 Mus muscu
5	20.8	83.2	202972	10	AL663052	AL663052 Mouse DNA
6	20.4	81.6	166829	2	AC146500	AC146500 Danio rer
7	20.2	80.8	178491	2	AL954819	AL954819 Rattus rer
8	20.2	80.8	222745	2	AC127100	AC127100 Rattus no
9	19.8	79.2	1019	6	AX655462	AX655462 Sequence
10	19.8	79.2	2993	8	AF391109	AF391109 Oryza sat
11	19.8	79.2	26845	2	AC116314	AC116314 Trypanoso
12	19.8	79.2	86392	2	AC135183	AC135183 Homo sapi
13	19.8	79.2	100849	9	HUNNEUROF	LO3367 Human oligo
14	19.8	79.2	113200	9	AC134669	AC134669 Homo sapi
15	19.8	79.2	139503	2	AP005289	AP005289 Oryza sat
16	19.8	79.2	160515	9	AC135724	AC135724 Homo sapi
17	19.8	79.2	163543	2	AP005072	AP005072 Oryza sat
18	19.8	79.2	187234	2	AC114142	AC114142 Rattus no
19	19.8	79.2	218347	2	AC126889	AC126889 Rattus no
20	19.8	79.2	297898	9	AC004526	AC004526 Homo sapi
21	19.4	77.6	175139	2	AC124103	AC124103 Mus muscu
22	19.4	77.6	209907	10	AC118038	AC118038 Mus muscu
23	19.2	76.8	70313	2	AC016216	AC016216 Homo sapi
24	19.2	76.8	86113	9	AF285442	AF285442 Homo sapi
25	19.2	76.8	101281	2	AC026389	AC026389 Homo sapi
26	19.2	76.8	104017	9	AC105050	AC105050 Homo sapi
27	19.2	76.8	107430	9	AC006988	AC006988 Homo sapi
28	19.2	76.8	112527	9	AC012356	AC012356 Homo sapi
29	19.2	76.8	132775	2	AC046156	AC046156 Homo sapi
30	19.2	76.8	137924	2	AC026684	AC026684 Homo sapi
31	19.2	76.8	142959	2	AC025453	AC025453 Homo sapi
32	19.2	76.8	144364	2	AP005876	AP005876 Oryza sat
33	19.2	76.8	158310	2	EX294108	EX294108 Mus muscu
34	19.2	76.8	164376	2	AC010823	AC010823 Homo sapi
35	19.2	76.8	164550	9	CNS01RHY	AL182683 Human chr
36	19.2	76.8	164603	9	AC025284	AC025284 Homo sapi
37	19.2	76.8	169199	2	AC109292	AC109292 Mus muscu
38	19.2	76.8	169883	2	AC023407	AC023407 Homo sapi
39	19.2	76.8	173829	10	AC122415	AC122415 Mus muscu
40	19.2	76.8	175208	10	AC112156	AC112156 Mus muscu
41	19.2	76.8	175577	2	AC022462	AC022462 Homo sapi
42	19.2	76.8	175749	9	AC025857	AC025857 Homo sapi
43	19.2	76.8	175988	9	CNS05TC7	AL355073 Human chr
44	19.2	76.8	176267	9	AC113420	AC113420 Homo sapi
45	19.2	76.8	181087	10	AC125080	AC125080 Mus muscu

ALIGNMENTS

RESULT 1	MMFHEAD1	6021 bp	DNA	linear	ROD 14-MAY-1997
LOCUS	M.musculus MFH-1 gene.				
DEFINITION	Y08222				
ACCESSION	Y08222.1	GI:1869968			
VERSION	mesenchyme fork head-1 protein; MFH-1 gene.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				

REFERENCE  
1  
Miura,N., Iida,K., Kakinuma,H., Yang,X.L. and Sugiyama,T.  
Isolation of the mouse (MFH-1) and human (FKHL 14) mesenchyme fork  
head-1 genes reveals conservation of their gene and protein

```

structures
JOURNAL Genomics 41 (3), 489-492 (1997)
MEDLINE 97312712
PUBMED 9163153
REFERENCE 2 (bases 1 to 6021)
AUTHORS Miura,N.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1996) N. Miura, Akita University School of
Medicine, Department of Biochemistry, 1-1-1 Hondo, Akita 010, JAPAN

FEATURES
source
1..6021
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129"
/db_xref="taxon:10090"
1486..4345
/feature="MFI-1"
2070..3554
/feature="MFI-1"
2070..3554
/codon_start=1
/product="mesenchyme fork head-1 protein"
/protein_id="CAA69399.1"
/db_xref="GI:1869969"
/db_xref="GOA:Q61850"
/db_xref="SWISS-PROT:Q61850"
/translation="MQARYSVSDNALGVPLVSEQNYVRAAGSYGGMASPMGVYSGH
PEQYAGNGRYSAPYHHQPAAPKDLVPPYVYIALITMAIONAPEKKITLNGIYQIM
DRFFPRENKQWNSIRHNLSECFVPRDRDKGKSGSYWTLDPDPSYNNPENGSP
LRRRRFKKDKPKDEARHLKEFPSTTAKGPTGTVPADGPKAEKPKVYVKEAAS
PALVITKVELSPGALQASPRASSTPAGSDGSLPEHHAAPNGLPGFVSVEITMT
LPTSPGDLSPARACLVVPLALPYAAPAAVTPCAOGLAAGSGAGYVCSWRA
MSLYTGERPAHVCPVALDEALSDHSGPSPLGALNLAAGOGALGASGHHOHGG
HLHQAPPPAPPPPPQAPATQATSWYLNHGDLHLPGHTFATQQQTFFNVREMS
HRLGDNSSIGESQVSNACOLPYRATPSLYRAAPYSYDCTKI"
4320..4325

polyA_signal
ORIGIN
Query Match 100.0%; Score 25; DB 10; Length 6021;
Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACTGTCC 25
Db 358 CAGTCCACAAATAAACAACTGTCC 382

RESULT 2
AC009108/c
LOCUS AC009108 168656 bp DNA linear PRI 29-OCT-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-46309, complete sequence.
ACCESSION AC009108
VERSION AC009108.10 GI:24418066
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168656)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168656)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 168656)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell

```

```

Drive, Walnut Creek, CA 94598, USA
On Oct 29, 2002 this sequence version replaced gi:13786306.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www-sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.2.
Location/Qualifiers
1..168656
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-46309"

ORIGIN
Query Match 100.0%; Score 25; DB 9; Length 168656;
Best Local Similarity 100.0%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTCCACAAATAAACAACTGTCC 25
Db 78355 CAGTCCACAAATAAACAACTGTCC 78331

RESULT 3
AC127554 178416 bp DNA linear ROD 27-NOV-2003
LOCUS AC127554
DEFINITION Mus musculus BAC clone RP24-323K23 from chromosome 8, complete
sequence.
ACCESSION AC127554
VERSION AC127554.4 GI:33457241
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 178416)
AUTHORS Cordes M. and Haglund K.
TITLE The sequence of Mus musculus BAC clone RP24-323K23
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 178416)
AUTHORS Wilson R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 178416)
AUTHORS McPherson J.D. and Waterston R.H.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 178416)
AUTHORS Wilson R.K.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 178416)
AUTHORS Wilson R.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 178416)
AUTHORS Wilson R.K.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 6, 2003 this sequence version replaced gi:33342444.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu

```

----- Summary Statistics  
 Center Project name: M\_B0323K23  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124170.

#### FEATURES

source

Location/Qualifiers

1..178416

/organism="Mus musculus"

/mol\_type="Genomic DNA"

/db\_xref="taxon:10090"

/chromosome="8"

/map="8"

/clone="RP24-323K23"

/clone\_lib="RPCI-24"

1791..1839

/rpt\_family="Alu"

3054..3288

/rpt\_family="B4"

3834..3845

/rpt\_family="B4"

3846..3972

/rpt\_family="Alu"

3973..4026

/rpt\_family="B4"

6269..8820

/note="CPG island (%GC=69.2, o/e=0.78, #CPGs=269)"

16848..16980

/rpt\_family="MIR"

18383..18494

/rpt\_family="Alu"

18974..19317

/rpt\_family="MaLR"

19821..20013

/rpt\_family="B2"

20188..20538

/rpt\_family="MaLR"

20539..20683

/rpt\_family="Alu"

20684..20799

/rpt\_family="MaLR"

22901..22957

/rpt\_family="ID"

22968..23090

/rpt\_family="Alu"

24384..24441

/rpt\_family="B4"

repeat\_region 25838..26022  
 /rpt\_family="MER1\_type"  
 repeat\_region 27267..27380  
 /rpt\_family="Alu"  
 repeat\_region 28700..28893  
 /rpt\_family="B2"  
 repeat\_region 29950..30049  
 /rpt\_family="B4"  
 repeat\_region 30809..30854  
 /rpt\_family="ID"  
 repeat\_region 30859..30963  
 /rpt\_family="Alu"  
 repeat\_region 30987..31062  
 /rpt\_family="ID"  
 repeat\_region 31157..31294  
 /rpt\_family="Alu"  
 repeat\_region 32192..32321  
 /rpt\_family="B4"  
 repeat\_region 32923..33036  
 /rpt\_family="Alu"  
 repeat\_region 33141..33245  
 /rpt\_family="Alu"  
 repeat\_region 33394..33521  
 /rpt\_family="B4"  
 repeat\_region 33806..33891  
 /rpt\_family="B4"  
 repeat\_region 34643..34976  
 /rpt\_family="MaLR"  
 repeat\_region 35547..35686  
 /rpt\_family="B2"  
 repeat\_region 35701..35839  
 /rpt\_family="Alu"  
 repeat\_region 35873..35982  
 /rpt\_family="Alu"  
 misc\_feature 37996..41776  
 /note="CPG island (%GC=69.3, o/e=0.84, #CPGs=225)"  
 repeat\_region 47196..47354  
 /rpt\_family="B2"  
 repeat\_region 47365..47498  
 /rpt\_family="Alu"  
 repeat\_region 47507..47550  
 /rpt\_family="B4"  
 repeat\_region 47706..47840  
 /rpt\_family="Alu"  
 repeat\_region 47716..47892  
 /rpt\_family="B4"  
 repeat\_region 47952..48067  
 /rpt\_family="Alu"  
 misc\_feature 51423..52857  
 /note="CPG island (%GC=66.5, o/e=0.72, #CPGs=121)"  
 repeat\_region 53367..53484  
 /rpt\_family="Alu"  
 repeat\_region 54128..54233  
 /rpt\_family="Alu"  
 repeat\_region 55219..55240  
 /rpt\_family="ERVK"  
 repeat\_region 55241..55364  
 /rpt\_family="Alu"  
 repeat\_region 55365..55460  
 /rpt\_family="ERVK"  
 repeat\_region 56094..56379  
 /rpt\_family="B4"  
 repeat\_region 59534..59702  
 /rpt\_family="B2"  
 repeat\_region 60089..60165  
 /rpt\_family="ID"  
 repeat\_region 60513..60709  
 /rpt\_family="B2"  
 repeat\_region 60886..61042  
 /rpt\_family="B4"  
 repeat\_region 61852..61997  
 /rpt\_family="MER1\_type"  
 repeat\_region 62247..62549